Space-Time Covariance Structures and Models

Wanfang Chen[1], Marc G. Genton[1], Ying Sun[1]

[1]Statistics Program, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, 23955-6900; email: wanfang.chen@kaust.edu.sa, marc.genton@kaust.edu.sa, ying.sun@kaust.edu.sa

Xxxx. Xxx. Xxx. Xxx. 2019. AA:1–26

https://doi.org/10.1146/((please add article doi))

Keywords

asymmetry, full symmetry, kriging, positive definiteness, random field, separability, stationarity

Abstract

In recent years, there has been a growing interest in modeling spatio-temporal data generated from monitoring networks, satellite imaging and climate models. Under Gaussianity, the covariance function is core to spatio-temporal modeling, inference and prediction. In this article, we review the various space-time covariance structures in which simplified assumptions, such as separability and full symmetry, are made to facilitate computation, and associated tests intended to validate these structures. We also review recent developments on constructing space-time covariance models, which can be separable or non-separable, fully symmetric or asymmetric, stationary or non-stationary, univariate or multivariate, and in Euclidean spaces or on the sphere. We visualize some of the structures and models with visuanimations. Finally, we discuss inference for fitting space-time covariance models, and describe a case study based on a new wind-speed dataset.

1. INTRODUCTION

Spatio-temporal variability is central to statistical analyses for dynamic processes in environmental science and geophysics. In geostatistics, the observations are typically modeled by a Gaussian random process, and the space-time covariance function plays an important role in describing dependence in the data and predicting realizations at unobserved sites (also called kriging) or future time points. Much progress has been made since the reviews of Kyriakidis & Journel (1999), Christakos (2000), Gneiting & Schlather (2002), Kolovos et al. (2004) and Gneiting et al. (2007). In this work, we provide a comprehensive review of recent advances in the literature on space-time covariance structures and models. A case study for wind-speed data illustrates the process of assessing the underlying space-time covariance structure, choosing appropriate models and performing inference.

Consider a space-time random process $Z(\mathbf{s}, t)$, where $\mathbf{s} \in \mathbb{R}^d$ $(d \geq 1)$ denotes a spatial location and $t \in \mathbb{R}$ denotes a time point. Henceforth, we assume that second moments of $Z(\mathbf{s}, t)$ exist and are finite. Denote the N space-time coordinates as $(\mathbf{s}_1, t_1), \ldots, (\mathbf{s}_N, t_N) \in \mathbb{R}^d \times \mathbb{R}$, and the covariance function of $Z(\mathbf{s}, t)$ by $\mathrm{C}(\mathbf{s}_1, t_1, \mathbf{s}_2, t_2) = \mathrm{cov}\{Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)\}$, where $(\mathbf{s}_1, t_1)$ and $(\mathbf{s}_2, t_2)$ in $\mathbb{R}^d \times \mathbb{R}$ are space-time coordinates.

A necessary and sufficient condition for a function C to be a covariance is positive definiteness (Yaglom 1987; De Iaco et al. 2011): C should satisfy

$$\sum_{i,j=1}^{N} c_i c_j \mathrm{C}(\mathbf{s}_i, t_i, \mathbf{s}_j, t_j) \geq 0, \qquad 1.$$

for all finite N, all $(\mathbf{s}_1, t_1), \ldots, (\mathbf{s}_N, t_N) \in \mathbb{R}^d \times \mathbb{R}$ and all real $c_1, \ldots, c_N$. This means that the covariance matrix of the random vector $\{Z(\mathbf{s}_1, t_1), \ldots, Z(\mathbf{s}_N, t_N)\}^\top$, denoted by $\boldsymbol{\Sigma}$, whose (i, j)th entry is $\sigma_{ij} = \mathrm{C}(\mathbf{s}_i, t_i, \mathbf{s}_j, t_j)$, is non-negative definite for all finite collections of coordinates $(\mathbf{s}_1, t_1), \ldots, (\mathbf{s}_N, t_N) \in \mathbb{R}^d \times \mathbb{R}$. The covariance function C is strictly positive definite if the left-hand side of Equation 1 is positive for any choice of distinct coordinates and any non-zero vector $(c_1, \ldots, c_N)^\top$. Strict positive definiteness ensures the existence of a unique solution of the kriging system for spatial or spatio-temporal prediction.

A main benefit of stationarity is the pooling of pairs of measurements separated by the same space-time vector for use as a set of repetitions, which are never available in practice, for statistical inference. Informally, stationarity means that certain statistical properties do not change over space and/or time. In this paper, we discuss covariance, or weak, stationarity for stochastic processes. Specifically, the covariance function C is said to be spatially stationary if it depends on space only through the spatial lag, and temporally stationary if it depends on time only through the temporal lag. It is said to be stationary if it is both spatially and temporally stationary, where there exists a permissible (in particular, positive definite) covariance function C defined on $\mathbb{R}^d \times \mathbb{R}$ such that $\mathrm{cov}\{Z(\mathbf{s}_1, t_1), Z(\mathbf{s}_2, t_2)\} = \mathrm{C}(\mathbf{h}, u)$ for all $(\mathbf{h}, u)$ in $\mathbb{R}^d \times \mathbb{R}$, where $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2$ and $u = t_1 - t_2$. The restrictions $\mathrm{C}(\cdot, 0)$ and $\mathrm{C}(\mathbf{0}, \cdot)$ are purely spatial and purely temporal covariance functions, respectively. A related notion is that of spatial isotropy, where $\mathrm{C}(\mathbf{h}, u)$ is restricted to be a function depending on the Euclidean norm of $\mathbf{h}$ only.

A useful characterization of stationary space-time covariance functions is the spectral representation, using Bochner's theorem (Bochner 1955).

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

$$\mathrm{C}(\mathbf{h}, u) = \int\int e^{\iota(\mathbf{h}^\top \boldsymbol{\omega} + u\tau)} \mathrm{d}F(\boldsymbol{\omega}, \tau), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \qquad 2.$$

where $\iota = \sqrt{-1}$, and F is a finite, non-negative and symmetric measure on $\mathbb{R}^d \times \mathbb{R}$.

The function F in Equation 2 is called the spectral measure. If C is integrable, the corresponding density with respect to the Lebesgue measure is

$$f(\boldsymbol{\omega}, \tau) = (2\pi)^{-(d+1)} \int\int e^{-\iota(\mathbf{h}^\top \boldsymbol{\omega} + u\tau)} \mathrm{C}(\mathbf{h}, u) \mathrm{d}\mathbf{h}\mathrm{d}u, \quad (\boldsymbol{\omega}, \tau) \in \mathbb{R}^d \times \mathbb{R},$$

and is called the spectral density. If f exists, then $\mathrm{d}F(\boldsymbol{\omega}, \tau) = f(\boldsymbol{\omega}, \tau)\mathrm{d}\boldsymbol{\omega}\mathrm{d}\tau$ in Equation 2.

In practice, a nugget effect (i.e., a discontinuity at the origin) is often included in the fitted stationary space-time covariance functions. In the spatio-temporal context, the nugget effect could be purely spatial, purely temporal or spatio-temporal, and takes the general form

$$\mathrm{C}(\mathbf{h}, u) = \delta_1 I\{(\mathbf{h}, u) = (\mathbf{0}, 0)\} + \delta_2 I\{\mathbf{h} = \mathbf{0}\} + \delta_3 I\{u = 0\}, \ (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R},$$

where δ_1, δ_2 and δ_3 are non-negative constants, and I is an indicator function.

2. SPACE-TIME COVARIANCE STRUCTURES

2.1. Structures

The space-time covariance function $\mathrm{C}(\mathbf{s}_1, t_1, \mathbf{s}_2, t_2)$ generally depends on the space-time coordinates $(\mathbf{s}_1, t_1)$ and $(\mathbf{s}_2, t_2)$. However, due to the large size of space-time data, it is usually computationally expensive and sometimes infeasible to implement traditional techniques, such as kriging, based on the general form of C. Simplifying structures are often imposed on C, such as stationarity and spatial isotropy, as introduced in Section 1, and full symmetry and separability, as introduced below. The relationships among several space-time covariance structures were illustrated by Gneiting et al. (2007).

2.1.1. Separability. One of the key difficulties of constructing space-time covariance models is that it is in general nontrivial to check whether a function is positive definite. One simple solution is to rely on the separability assumption:

$$\mathrm{C}(\mathbf{s}_1, t_1, \mathbf{s}_2, t_2) = \mathrm{C}_\mathrm{S}(\mathbf{s}_1, \mathbf{s}_2) \cdot \mathrm{C}_\mathrm{T}(t_1, t_2), \qquad 3.$$

for all $(\mathbf{s}_1, t_1)$ and $(\mathbf{s}_2, t_2)$ in $\mathbb{R}^d \times \mathbb{R}$, where C_S and C_T are purely spatial and purely temporal covariance functions, respectively. The separability of C eases the computation, because the space-time covariance matrix factorizes into the Kronecker product of purely spatial and purely temporal covariance matrices. However, Equation 3 assumes no interaction between space and time, which is often unrealistic for physical processes.

2.1.2. Stationarity and separability. A stationary space-time covariance function C is separable if $\mathrm{C}(\mathbf{h}, u) = \mathrm{C}_\mathrm{S}(\mathbf{h}) \cdot \mathrm{C}_\mathrm{T}(u)$ for all $(\mathbf{h}, u)$ in $\mathbb{R}^d \times \mathbb{R}$, where C_S and C_T are stationary purely spatial and purely temporal covariance functions, respectively. Equivalently, the covariance function C is separable if it can be factorized in terms of the joint process as

$$\mathrm{C}(\mathbf{h}, u) = \mathrm{C}(\mathbf{h}, 0) \cdot \mathrm{C}(\mathbf{0}, u)/\mathrm{C}(\mathbf{0}, 0), \qquad 4.$$

for all $(\mathbf{h}, u)$ in $\mathbb{R}^d \times \mathbb{R}$ (Mitchell et al. 2005). In the spectral domain, a stationary covariance function is separable if and only if the spectral measure factors as a product of measures over the spatial and temporal domains. In particular, if the spectral density exists, it factors as a product over the domains.

2.1.3. Full symmetry. The space-time covariance function C is fully symmetric if $\mathrm{C}(\mathbf{s}_1, t_1, \mathbf{s}_2, t_2) = \mathrm{C}(\mathbf{s}_1, t_2, \mathbf{s}_2, t_1)$ for all $(\mathbf{s}_1, t_1)$ and $(\mathbf{s}_2, t_2)$ in $\mathbb{R}^d \times \mathbb{R}$. Full symmetry contains separability as a special case, so covariance structures that are not fully symmetric are non-separable. There is often a lack of full symmetry for processes in geoscience, meteorology, hydrology and ecology due to transport effects such as prevailing air or water flows and atmosphere circulation.

2.1.4. Stationarity and full symmetry. A stationary space-time covariance function C is fully symmetric if

$$\mathrm{C}(\mathbf{h}, u) = \mathrm{C}(\mathbf{h}, -u) = \mathrm{C}(-\mathbf{h}, u) = \mathrm{C}(-\mathbf{h}, -u), \qquad 5.$$

for all $(\mathbf{h}, u)$ in $\mathbb{R}^d \times \mathbb{R}$. In the purely spatial context, this property is also known as axial symmetry (Scaccia & Martin 2005) or reflection symmetry (Lu & Zimmerman 2005b). A stationary and fully symmetric space-time covariance function can also be characterized in terms of the spectral measure or spectral density by Bochner's theorem; see Theorem 4.3.2 and discussion in Gneiting et al. (2007).

2.1.5. Stationarity and Taylor's hypothesis. A stationary space-time covariance function C satisfies Taylor's hypothesis (Taylor 1938) if there exists a velocity vector $\mathbf{v} \in \mathbb{R}^d$ such that $\mathrm{C}(\mathbf{0}, u) = \mathrm{C}(\mathbf{v}u, 0),\ u \in \mathbb{R}$. Qualitatively, this implies that the process evolves slowly in time relative to the advective time scale, and it is often referred to as the frozen field hypothesis. As a specific asymmetric covariance structure, Taylor's hypothesis has gained wide interest in meteorology and hydrology, and has been found to be plausible for various space-time covariance models (Gupta & Waymire 1987; Cox & Isham 1988). Examples of covariance functions that admit Taylor's hypothesis exactly can be found in Gneiting et al. (2007).

2.2. Testing and Visualizing the Structure

The aforementioned various space-time covariance structures of stationarity, spatial isotropy, separability, full symmetry and Taylor's hypothesis are desirable from a computational point of view. However, these assumptions are not always realistic; see, e.g., the graphical evidence in Cressie & Huang (1999), Gneiting et al. (2007) and Jun & Stein (2007). Formal testing procedures have been developed to restrict the class of models one needs to investigate when modeling space-time data.

2.2.1. Tests for stationarity. Tests for spatial stationarity have been proposed by Fuentes (2005), who extended a test for stationarity for time series to spatial random fields, and Bandyopadhyay & Rao (2017), who used a spectral method for testing stationarity for irregularly spaced spatial data, based on observing that the Fourier transform of a stochastic process is nearly uncorrelated if the process is stationary but correlated if it is non-stationary. Tests for spatio-temporal stationarity include those in Jun & Genton (2012), which are based on the asymptotic normality of the proposed test statistics that are functions of estimators of covariances at selected spatial and temporal lags under spatial stationarity, and Bandyopadhyay et al. (2017), who extended the method in Bandyopadhyay & Rao (2017) to space-time data over discrete time and irregular spatial locations. A review of tests for spatial isotropy can be found in Weller & Hoeting (2016).

2.2.2. Tests for separability. Early tests for separability of spatio-temporal processes were based on parametric models (Guo & Billard 1998; Brown et al. 2001). Later researchers have focused on likelihood ratio tests (Roy & Khattree 2003, 2005a,b; Roy & Leiva 2008; Lu & Zimmerman 2005a; Mitchell et al. 2005, 2006; Simpson 2010), and tests based on properties of the spectral representation (Scaccia & Martin 2002, 2005; Fuentes 2006; Crujeiras et al. 2010). Likelihood ratio tests were developed in the context of multivariate repeated measures, and Mitchell et al. (2005) adapted the test of Mitchell et al. (2006) to the spatio-temporal setting. Bevilacqua et al. (2010) presented parametric tests based on weighted composite likelihood estimates in situations where classical methods such as a likelihood ratio test may fail. Aston et al. (2017), Liu et al. (2017) and Constantinou et al. (2017) constructed nonparametric tests for separability of space-time functional processes. Cappello et al. (2018) proposed a method for testing positive and negative non-separabilities, whose definitions are given in Section 3.1.3.

2.2.3. Tests for symmetry. For testing lack of symmetry for spatial processes, Scaccia & Martin (2005) used spectral methods for assessing axial symmetry, and Lu & Zimmerman (2005b) used two-dimensional spatial periodograms for assessing axial symmetry and complete symmetry. For testing full symmetry in a spatio-temporal setting, the nonparametric test of Park & Fuentes (2008) is based on the spectral representation of spatio-temporal processes. Since full symmetry implies separability, covariance structures that are not fully symmetric are non-separable, and hence tests for full symmetry can be used to reject separability. In terms of testing the Taylor's hypothesis, Li et al. (2009) proposed a formal statistical testing procedure based on the asymptotic joint normality of covariance estimators derived by Li et al. (2008a).

2.2.4. Unified tests for separability and symmetry. Li et al. (2007) proposed a unified nonparametric methodology for testing various assumptions (i.e., separability, full symmetry, isotropy and Taylor's hypothesis) based on the asymptotic joint normality of sample space-time covariance estimators for a selected spatio-temporal lag set, and Li et al. (2008b) extended the test to the multivariate space-time setting. Shao & Li (2009) modified the tests in Li et al. (2007, 2008b) by employing an inconsistent estimator of the asymptotic covariance matrix, with the advantage that the test is free of any tuning parameters.

Huang & Sun (2019) proposed test functions for separability and full symmetry,

$$f_{\mathbf{h}}(u) = \mathrm{C}(\mathbf{h},u)/\mathrm{C}(\mathbf{h},0) - \mathrm{C}(\mathbf{0},u)/\mathrm{C}(\mathbf{0},0), \quad g_{\mathbf{h}}(u) = \mathrm{C}(\mathbf{h},u) - \mathrm{C}(\mathbf{h},-u), \qquad 6.$$

respectively, as functions of the temporal lag u for any spatial lag $\mathbf{h}$. The function $f_{\mathbf{h}}(u)$ ($g_{\mathbf{h}}(u)$) is 0 for any u and $\mathbf{h}$ if C is separable (fully symmetric); otherwise, the test function will move away from zero. They then used functional boxplots (Sun & Genton 2011) to visualize the functional median and variability of the test functions. The degree of non-separability or asymmetry is indicated by the extent of departure from zero at all temporal lags in the functional boxplot. The authors also developed a rank-based nonparametric testing procedure to assess the significance of the non-separability or asymmetry.

3. SPACE-TIME COVARIANCE MODELS

Owing to their computational convenience, separable space-time covariance functions have been applied to multivariate repeated measures or doubly repeated measures (Chaganty

& Naik 2002) and to spatio-temporal models (Huizenga et al. 2002; Mitchell & Gumpertz 2003). Although separability is valuable, it does not allow for space-time interaction in the covariance. Stein (2005a) pointed out that separable models are not smoother away from the origin than they are at the origin, leading to a kind of discontinuity in certain correlations that one might wish to avoid in some circumstances. More detailed discussion of the shortcomings of separable models can be found in Kyriakidis & Journel (1999) or Cressie & Huang (1999).

Recent studies have focused on constructing non-separable models, which are often physically more realistic. The associated computational burden can be reduced in some settings by using the separable approximation of the non-separable space-time covariance matrix proposed by Genton (2007). Other techniques that allow for fast inference on large space-time data sets are discussed in Section 4.

Non-separable space-time covariance models can be constructed from Fourier transforms of permissible spectral densities, mixtures of separable models, and partial differential equations representing physical laws. They can be fully symmetric or asymmetric, stationary or non-stationary, univariate or multivariate, and in the Euclidean space or on the sphere. We discuss these models in detail below.

3.1. Stationary and Fully Symmetric Covariance Models (SSCMs)

3.1.1. SSCMs based on spectral methods. SSCMs can be constructed through Fourier transforms of permissible spectral densities, based on the spectral representation, Equation 2, of covariance functions. A class of SSCMs has been characterized by Cressie & Huang (1999) through a closed-form Fourier inversion in $\mathbb{R}^d$,

$$\mathrm{C}(\mathbf{h}, u) = \int e^{\iota \mathbf{h}^\top \boldsymbol{\omega}} \rho(\boldsymbol{\omega}, u) \mathrm{d}\boldsymbol{\omega}, \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R},$$

where $\rho(\boldsymbol{\omega}, u), u \in \mathbb{R}$, is a continuous positive-definite function for all $\boldsymbol{\omega} \in \mathbb{R}^d$. This approach is restricted to a small class of functions for which a closed-form solution to a d-variate Fourier integral is known.

Gneiting (2002) proposed a more general class of SSCMs using the approach of Cressie & Huang (1999) but without the Fourier integral limitation,

$$\mathrm{C}(\mathbf{h}, u) = \frac{\sigma^2}{\psi(u^2)^{d/2}} \varphi \left\{ \frac{\|\mathbf{h}\|^2}{\psi(u^2)} \right\}, \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \qquad 7.$$

where $\sigma^2 = \mathrm{C}(\mathbf{0}, 0)$, $\varphi(r) \geq 0, r \geq 0$, is a completely monotone function (i.e., it possesses derivatives of all orders with alternating signs), and $\psi(r) > 0, r \geq 0$, has a completely monotone derivative (i.e., ψ is a Bernstein function).

Porcu et al. (2006) generalized the Gneiting (2002) class of covariance models to $\mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \times \mathbb{R}^{d_3} \times \mathbb{R}^{d_4}$, where $d_i \in \mathbb{N}$, $i = 1, 2, 3, 4$, and built spatially component-wise anisotropic and temporally symmetric SSCMs defined on $\mathbb{R}^3 \times \mathbb{R}$,

$$\mathrm{C}(h_1, h_2, h_3, u) = \frac{\sigma^2}{\psi_1(|h_1|^2)^{1/2} \psi_2(|u|^2)^{1/2}} \times \mathcal{L} \left\{ \frac{|h_2|^2}{\psi_1(|h_1|^2)}, \frac{|h_3|^2}{\psi_2(|u|^2)} \right\}, \quad h_1, h_2, h_3, u \in \mathbb{R}, \qquad 8.$$

where $\sigma^2 = \mathrm{C}(0, 0, 0, 0)$, ψ_1 and ψ_2 are either Bernstein functions, variograms or increasing and concave functions on $[0, +\infty)$, and $\mathcal{L}$ is defined as before. A similar construction was proposed by Porcu et al. (2007): $\mathrm{C}(\mathbf{h},u) = \mathcal{L}\big\{\sum_{i=1}^{d}\psi_i(|h_i|),\psi_t(|u|)\big\}, (\mathbf{h},u) \in \mathbb{R}^d \times \mathbb{R}$, where $\mathbf{h} = (h_1,\dots,h_d)^\top$, and $\mathcal{L}$ and ψ are defined as above.

Stein (2005a) showed that similar to separable models, the SSCMs proposed by Cressie & Huang (1999), Gneiting (2002) and De Iaco et al. (2001, 2002) (to be discussed in Section 3.1.2) are not differentiable at the origin or along certain axes, which leads to discontinuities of the autocorrelation function away from the origin. He proposed a parametric class of spectral densities of the form

$$f(\boldsymbol{\omega},\tau) = \{P_1(\|\boldsymbol{\omega}\|^2) + P_2(|\tau|^2)\}^{-\nu}, \quad (\boldsymbol{\omega},\tau) \in \mathbb{R}^d \times \mathbb{R},$$

where $\nu > 0$, f is bounded, P_1 and P_2 are non-negative polynomials on $[0,+\infty)$ of positive degrees α_1 and α_2, respectively, and $d/(\alpha_1\nu) + 1/(\alpha_2\nu) < 2$. The corresponding space-time covariance function C can achieve any degree of differentiability in space and in time away from the origin. Stein (2005a) also described a general approach to constructing space-time covariance functions that are spatially isotropic but asymmetric from symmetric models by taking derivatives.

Porcu et al. (2008) built spectral densities whose Fourier pair is infinitely differentiable away from the origin. The new spectral density is a product of two spectral densities, f_1 and f_2, under the condition of integrability,

$$f(\boldsymbol{\omega},\tau) = f_1(a_1|\tau|^\alpha + b_1\|\boldsymbol{\omega}\|^\beta)f_2(a_2|\tau|^\alpha + b_2\|\boldsymbol{\omega}\|^\beta), \quad (\boldsymbol{\omega},\tau) \in \mathbb{R}^d \times \mathbb{R},$$

where a_1, a_2, b_1, b_2 are positive constants, and $\alpha, \beta \in \mathbb{N}$ are even numbers. This general class includes both the separable and non-separable cases in the spectral domain.

The Matérn class (Matérn 1986) is an important family of covariance functions in spatial statistics. Its spectral density is $f(\boldsymbol{\omega}) = \gamma(\alpha^2 + \|\boldsymbol{\omega}\|^2)^{-\nu-d/2}, \boldsymbol{\omega} \in \mathbb{R}^d$, where $\gamma > 0$, $\alpha > 0$ and $\nu > (d+1)/2$ are the scale, spatial range and smoothness parameters, respectively. Fuentes et al. (2008) extended the Matérn class to the spatio-temporal setting, where the spectral density has the form

$$f(\boldsymbol{\omega},\tau) = \gamma(\alpha^2\beta^2 + \beta^2\|\boldsymbol{\omega}\|^2 + \alpha^2\tau^2 + \epsilon\|\boldsymbol{\omega}\|^2\tau^2)^{-\nu}, \quad (\boldsymbol{\omega},\tau) \in \mathbb{R}^d \times \mathbb{R}, \qquad 9.$$

where β is the temporal range parameter, and $\epsilon \in [0,1]$ is the non-separability parameter. If $\epsilon = 1$, the corresponding covariance is separable, and both the spatial and temporal components are Matérn-type covariances. If $\epsilon \neq 1$ and $\nu < \infty$, the corresponding covariance is non-separable. Ip & Li (2017b) provided closed forms of the covariance functions for $\epsilon = 1$ and $\epsilon = 0$; the closed form does not exist in general when $\epsilon \in (0,1)$, in which case one must rely on a numerical Fourier transformation.

Horrell & Stein (2017) defined a half-spectral representation either in space if only the integration over $\boldsymbol{\omega}$ is performed in Equation 2, or in time if only the integration over τ is performed. The models in Cressie & Huang (1999) and Gneiting (2002) are half-spectral in space, and that in Stein (2005b) is half-spectral in time. Horrell & Stein (2017) also developed a class of SSCMs that is half-spectral in time.

3.1.2. SSCMs based on mixtures. SSCMs can also be constructed as mixtures of separable covariances. As a result, the mixture models are fully symmetric.

De Iaco et al. (2001) proposed a simple product-sum SSCM,

$$\mathrm{C}(\mathbf{h},u) = a_0\mathrm{C}_{\mathrm{S}0}(\mathbf{h})\mathrm{C}_{\mathrm{T}0}(u) + a_1\mathrm{C}_{\mathrm{S}1}(\mathbf{h}) + a_2\mathrm{C}_{\mathrm{T}2}(u), \quad (\mathbf{h},u) \in \mathbb{R}^d \times \mathbb{R}, \qquad 10.$$

where a_0, a_1 and a_2 are non-negative constants, and $\mathrm{C}_{\mathrm{S}0}, \mathrm{C}_{\mathrm{S}1}$ and $\mathrm{C}_{\mathrm{T}0}, \mathrm{C}_{\mathrm{T}2}$ are stationary, purely spatial and purely temporal covariance functions, respectively. De Iaco et al. (2002) extended Equation 10 as

$$\mathrm{C}(\mathbf{h}, u) = \int \{a_0 \mathrm{C}_\mathrm{S}(\mathbf{h}; \theta)\mathrm{C}_\mathrm{T}(u; \theta) + a_1 \mathrm{C}_\mathrm{S}(\mathbf{h}; \theta) + a_2 \mathrm{C}_\mathrm{T}(u; \theta)\}\mathrm{d}\mu(\theta), \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R},$$

under integrability conditions, where μ is a finite, non-negative measure on a non-empty set Θ, $\mathrm{C}_\mathrm{S}(\cdot; \theta)$ and $\mathrm{C}_\mathrm{T}(\cdot; \theta)$ are stationary, purely spatial and purely temporal covariance functions, respectively, for each $\theta \in \Theta$, and a_0, a_1 and a_2 are non-negative constants. When $a_1 = a_2 = 0$, this corresponds to an extension of the Cressie & Huang (1999) model.

Ma (2002) constructed SSCMs by introducing a non-negative bivariate discrete random vector $(U, V)^\top$, conditional on which the underlying space-time random process Z possesses a separable covariance structure. Then the (unconditional) correlation function of Z, defined as $\rho(\mathbf{h}, u) = \mathrm{C}(\mathbf{h}, u)/\mathrm{C}(\mathbf{0}, 0)$, has the form

$$\rho(\mathbf{h}, u) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \rho_\mathrm{S}(\mathbf{h}; i)\rho_\mathrm{T}(u; j)p_{ij}, \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \qquad 11.$$

where $\{p_{ij}, (i, j) \in \mathbb{Z}_+^2\}$ is the probability mass function of $(U, V)^\top$, and conditional on $(U, V) = (i, j) \in \mathbb{Z}_+^2$, $\rho_\mathrm{S}(\cdot; i)$ and $\rho_\mathrm{T}(\cdot; j)$ are stationary, purely spatial and purely temporal correlation functions, respectively. A similar conditional method was used in Ma (2003a,b). Ma (2005) derived SSCMs via linear combinations of space-time covariance functions, which are formulated by using the cosine transform method in Ma (2003b).

Porcu et al. (2008) joined several approaches from De Iaco et al. (2001), Gneiting (2002) and Stein (2005a) to construct SSCMs while removing some undesirable features of the previously proposed models. However, the level of differentiability in their model cannot be determined analytically. This drawback is covered by their second proposal based on the spectral method, which was discussed in Section 3.1.1.

3.1.3. Measures of non-separability. Recent work on constructing valid SSCMs has focused on different types and degrees of non-separability that characterize the extent of space-time interaction, or the dependence between spatial and temporal components of the process.

Based on the characterization of separability in Equation 4, Rodrigues & Diggle (2010) proposed simple definitions for positive and negative non-separability of space-time covariance functions, corresponding to the ratio

$$\frac{\mathrm{C}(\mathbf{h}, u) \times \mathrm{C}(\mathbf{0}, 0)}{\mathrm{C}(\mathbf{h}, 0) \times \mathrm{C}(\mathbf{0}, u)}$$

being larger or smaller than 1, respectively, for all $(\mathbf{h}, u)$ in $\mathbb{R}^d \times \mathbb{R}$. They showed that the SSCMs in Gneiting (2002) cannot accommodate negative non-separability, those in De Iaco et al. (2001) cannot accommodate positive non-separability, and those in Ma (2003a) accommodate non-separability but under the restriction that the temporal correlation structure decays exponentially. They then proposed SSCMs using a discretized convolution-based model with a non-separable kernel function, within which negative, zero or positive non-separability can be achieved by changing the value of a single parameter. The convolution-based model is shown to achieve substantial computational gains over the use of a Gaussian process with directly specified covariance structure when the process being modeled is spatially smooth, that is, when the correlation decays slowly with increasing spatial separation.

De Iaco & Posa (2013) pointed out that the definitions in Rodrigues & Diggle (2010) did not account for the fact that a covariance function usually depends on a vector of parameters $\theta \in \boldsymbol{\Theta}$, and proposed a definition of uniformly positive and negative non-separability, corresponding to the same relationships as in Rodrigues & Diggle (2010) but for all $(\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}$ and all $\theta \in \boldsymbol{\Theta}$. Thus, uniformly positive (negative) non-separability implies point-wise positive (negative) non-separability, but the converse is false. They showed that the covariance models in Gneiting (2002) and Ma (2002) are uniformly positive non-separable, and the model in De Iaco et al. (2001) is uniformly negative non-separable. They also proposed wide classes of SSCMs based on extensions of De Iaco et al. (2001), Ma (2002) and De Iaco et al. (2002), with changing non-separability indexes.

Fonseca & Steel (2011a) suggested a measure of non-separability for their specific proposed SSCM, formed by setting $U = X_0 + X_1$ and $V = X_0 + X_2$ in Ma (2002), where X_0, X_1 and X_2 are independent non-negative random variables with finite moment generating functions M_0, M_1 and M_2, respectively. Then the new SSCM is given by

$$\mathrm{C}(\mathbf{h}, u) = \sigma^2 M_0\{-\gamma_1(\mathbf{h}) - \gamma_2(u)\} M_1\{-\gamma_1(\mathbf{h})\} M_2\{-\gamma_2(u)\}, \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R},$$

where γ_1 and γ_2 are purely spatial and purely temporal variograms, respectively. This new SSCM allows for different degrees of smoothness across space and time and for long-range dependence in time. The measure of non-separability is the correlation between U and V as an indication of space-time interaction, which is always between 0 and 1, with 0 indicating separability and 1 meaning high dependence between space and time. However, this model excludes negative non-separability in the sense of Rodrigues & Diggle (2010).

Fonseca & Steel (2017) proposed a general measure of non-separability that can be applied to any class of non-separable covariance functions. Since negative non-separability is uncommon in practice, they only considered positively non-separable functions. They showed that some of the models proposed in the literature, such as those in Cressie & Huang (1999), Gneiting (2002) and Rodrigues & Diggle (2010), do not achieve strong non-separability and that the parameters do not always have clear interpretations. They also illustrated how their proposal can be applied to non-stationary processes.

3.2. Stationary and Asymmetric Covariance Models (SACMs)

3.2.1. Lagrangian reference frame. Atmospheric and geophysical processes are often asymmetric due to transport effects, such as prevailing air and water flows. To model such processes, we can use the Lagrangian reference frame (May & Julien 1998) to achieve asymmetry. Specifically, consider a spatial random field on $\mathbb{R}^d$ with a stationary covariance function $\mathrm{C_S}$, and suppose that the entire field moves time-forward with a random velocity vector $\mathbf{V} \in \mathbb{R}^d$. The resulting spatio-temporal random process has the covariance

$$\mathrm{C}(\mathbf{h}, u) = \mathrm{E}\{\mathrm{C_S}(\mathbf{h} - \mathbf{V}u)\}, \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}. \qquad 12.$$

Covariances of this form are generally asymmetric and satisfy Taylor's hypothesis. Ma (2003a), among others, proposed similar constructions, and established a theorem that formalizes the validity of turning a purely spatial covariance function into a space-time covariance function.

Various options can be physically justified for choosing the random velocity vector $\mathbf{V}$. For instance, in the context of atmospheric transport effects driven by prevailing winds, the simplest case occurs when $\mathbf{V} = \mathbf{v}$ is constant, which represents the mean wind vector as determined from synoptic or local wind patterns. Gupta & Waymire (1987) referred to this as the frozen field model. The non-frozen field model, proposed by Cox & Isham (1988), offers more flexibility by treating $\mathbf{V}$ as random. Finally, the distribution of $\mathbf{V}$ could be updated dynamically according to the current state of the atmosphere, which yields non-stationary, flow-dependent covariance structures.

The stationary frozen field model has been used to model winds (Gneiting et al. 2007; Ezzat et al. 2018), waves (Ailliot et al. 2011), solar irradiance (Inoue et al. 2012; Lonij et al. 2013; Shinozaki et al. 2016), and disease (Christakos 2017). However, it is isotropic for any $u \neq 0$. Stationary non-frozen field models, where the velocity vector is not constant, can introduce anisotropy. Porcu et al. (2006) proposed an anisotropic version of the model by partitioning $\mathbf{h}$ and $\mathbf{V}$ into smaller components,

$$\mathrm{C}(\mathbf{h}, u) = \mathrm{E}_{\mathbf{V}_1}\mathrm{E}_{\mathbf{V}_2}[\mathcal{L}\{\gamma_1(\mathbf{h}_1 - \mathbf{V}_1 u), \gamma_2(\mathbf{h}_2 - \mathbf{V}_2 u)\}], \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R},$$

where γ_1, γ_2 are purely spatial variograms, and $\mathcal{L}$ is the bivariate Laplace transform of a non-negative random vector. Schlather (2010) derived the explicit form of the non-frozen field model for specific choices of $\mathrm{C_S}$ and the distribution of $\mathbf{V}$.

3.2.2. Other physics-based SACMs. There is often a need for the covariance function to incorporate physical features of environmental variables. Covariance models that satisfy Taylor's hypothesis are also physically realistic, and the Lagrangian reference frame discussed above has its roots in physics. Many stationary and asymmetric space-time covariance functions can be derived as solutions of diffusion equations or stochastic partial differential equations. Related approaches have been discussed by Heine (1955), Jones & Zhang (1997), Christakos & Hristopulos (1998), Christakos (2000), Brown et al. (2000), Ma (2003a), and Kolovos et al. (2004).

3.3. Non-stationary Covariance Models (NSCMs)

There is an extensive literature on non-stationary spatial covariance models (e.g., Sampson & Guttorp 1992; Higdon et al. 1999; Fuentes & Smith 2001; Fuentes 2002; Nychka et al. 2002; Pintore & Holmes 2007). In this section, we focus on non-stationary space-time covariance models, which can be non-stationary in space and stationary in time, stationary in space and non-stationary in time, or non-stationary in both space and time.

Ma (2002) proposed non-separable NSCMs by scale mixtures in a similar way to the positive power mixtures model in Equation 11. Kolovos et al. (2004) obtained non-separable NSCMs from the generalized random field theory (Christakos 1991, 2004). Porcu et al. (2007) generalized the Gneiting (2002) class to spatial non-stationarity using non-stationary kernels. Fuentes et al. (2008) constructed non-separable NSCMs based on a mixture of local stationary spectral densities of the spatio-temporal Matérn-type as shown in Equation 9.

More sophisticated NSCMs have appeared recently. Ip & Li (2015) proposed covariance models that allow the model parameters to vary over time. Shand & Li (2017) formulated a spatio-temporal dimension expansion approach based on the work of Bornn et al. (2012). Xu & Gardoni (2018) constructed NSCMs based on the improved latent space approach.

3.4. Covariance Models on the Sphere (CMSPs)

Climate model outputs and remotely sensed networks have provided environmental, geophysical, and atmospheric sciences with many datasets that are inherently global and temporally evolving. Thus, related applications have increasingly focused on modeling processes evolving temporally over a sphere, using, for example, global data covering a large portion of the Earth. In principle, all the aforementioned methods in the large literature on covariance models for flat surfaces using the Euclidean distance can produce valid CMSPs using the chordal distance, since the chordal distance in $\mathcal{S}_r^{d-1}$, a sphere in $\mathbb{R}^d$ with radius r, is equivalent to the Euclidean distance in $\mathbb{R}^d$ (Jun & Stein 2007). The chordal distance, d_{CH}, is the length of the shortest straight line between two locations on the sphere.

However, the chordal distance is not a valid metric for the description of spatial dependence, since it does not respect the curvature of the sphere, and thus may result in physically unrealistic distortions (Gneiting 2013). In addition, it underestimates the true distance between the points on the sphere, which has a non-negligible impact on the estimation of the spatial scale (Porcu et al. 2016). Instead, the great circle (geodesic) distance, d_{GC}, defined as the shortest distance between any two points on a sphere measured along a path on its surface, is the physically most natural metric for processes on the surface of a sphere. Nevertheless, space-time covariances in Euclidean spaces are generally not valid on the sphere if coupled with the geodesic distance (Berg & Porcu 2017). For example, the Matérn model with the geodesic distance is valid on the sphere only under a severe restriction on the smoothing parameter (not smoother than the exponential covariance). This limitation is inherited for space-time processes, and it has motivated researchers to either use the chordal distance (North et al. 2011; Jeong & Jun 2015; Guinness & Fuentes 2016), or build new CMSPs, which might not be valid for the Euclidean distance, based on positive definite functions using the geodesic distance (Porcu et al. 2016; Berg & Porcu 2017; Alegría et al. 2019).

More sophisticated non-stationary spatial or spatio-temporal CMSPs have been proposed by Jun & Stein (2007, 2008), Hitczenko & Stein (2012) and Alegría & Porcu (2017), among others. Gneiting (2013) and Jeong et al. (2017) provided reviews on CMSPs in a spatial context. A thorough review on CMSPs on the sphere in $\mathbb{R}^3$ representing our planet has been given in Porcu et al. (2018), and Salvaña & Genton (2020) also included a review on spatial and spatio-temporal CMSPs.

3.5. Multivariate Cross-Covariance Models (MCCMs)

It is now common to monitor several geo-referenced variables evolving in time. Hence, it is important to specify an MCCM able to capture both the spatio-temporal and the cross-variable dependencies.

Consider a p-variate spatio-temporal random process $\mathbf{Z}(\mathbf{s},t) = \{Z_1(\mathbf{s},t),\ldots,Z_p(\mathbf{s},t)\}^\top$, $(\mathbf{s},t) \in \mathbb{R}^d \times \mathbb{R}$. The corresponding matrix-valued cross-covariance function at two spatio-temporal coordinates, $(\mathbf{s}_1,t_1),(\mathbf{s}_2,t_2) \in \mathbb{R}^d \times \mathbb{R}$, is $\mathbf{C}(\mathbf{s}_1,t_1,\mathbf{s}_2,t_2) = \{\mathrm{C}_{ij}(\mathbf{s}_1,t_1,\mathbf{s}_2,t_2)\}_{i,j=1}^p$, where $\mathrm{C}_{ij}(\mathbf{s}_1,t_1,\mathbf{s}_2,t_2) = \mathrm{cov}\{Z_i(\mathbf{s}_1,t_1),Z_j(\mathbf{s}_2,t_2)\}$, $i,j = 1,\ldots,p$. The cross-covariance matrix $\mathbf{\Sigma} \in \mathbb{R}^{Np\times Np}$ of the random vector $\{\mathbf{Z}(\mathbf{s}_1,t_1)^\top,\ldots,\mathbf{Z}(\mathbf{s}_N,t_N)^\top\}^\top \in \mathbb{R}^{Np}$, with $N \times N$ block elements of $p \times p$ matrices $\mathbf{C}(\mathbf{s}_k,t_k,\mathbf{s}_l,t_l)$, $k,l = 1,\ldots,N$, should satisfy the condition of non-negative definiteness, i.e., $\mathbf{a}^\top\mathbf{\Sigma}\mathbf{a} \geq 0$ for any integer N, any finite set of coordinates $(\mathbf{s}_1,t_1),\ldots,(\mathbf{s}_N,t_N)$, and any vector $\mathbf{a} \in \mathbb{R}^{Np}$.

The notions of stationarity and isotropy are defined in essentially the same way as in the univariate case. Arguably, the easiest way to build stationary MCCMs is again to rely on the assumption of (space-time) separability, where each $\mathrm{C}_{ij}(\mathbf{h},u)$ can be written as a product of a purely spatial cross-covariance $\mathrm{C}_{ij}(\mathbf{h})$ and purely temporal univariate covariance $\mathrm{C}_{\mathrm{T}}(u)$,

$(\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}$, and this implies full symmetry, which occurs when each C_{ij} satisfies Equation 5. Alegría et al. (2019) defined the notion of space-time p-separability and two related structures, and other notions of separability have been discussed by Apanasovich & Genton (2010) based on latent dimensions, and by De Iaco et al. (2019) based on the linear model of coregionalization (LMC). Again, separability assumptions can be useful for both modeling and estimation purposes since they alleviate the computational burdens, but are generally unrealistic. The literature has thus focused on developing non-separable models.

In the purely spatial context, Genton & Kleiber (2015) provided a review on MCCMs, including models built from univariate models based on LMC, convolution methods and latent dimensions, and the multivariate Matérn model (Gneiting et al. 2010). These models can be further adapted to the space-time context. For example, Bourotte et al. (2016) and Ip & Li (2016, 2017a) provided several adaptations of the multivariate purely spatial Matérn model to space-time, and Alegría et al. (2019) extended the modified Gneiting class introduced by Porcu et al. (2016) to the multivariate case, and also constructed an additional Gneiting type cross-covariance by adapting the latent dimension approach (Apanasovich & Genton 2010) to the spherical context. Several studies discussed in previous sections, for example, Rodrigues & Diggle (2010), Ip & Li (2015) and Porcu et al. (2016), also tackled the multivariate case. A more thorough review of MCCMs can be found in Salvaña & Genton (2020), who proposed a class of non-stationary Lagrangian MCCMs too. Spatio-temporal covariance models have also been discussed in Cressie & Wikle (2011), Montero et al. (2015), Christakos (2017) and Wikle et al. (2019).

3.6. Illustrations

In this section, we illustrate various space-time covariance structures and realizations through visuanimations (Genton et al. 2015). To illustrate non-separability, we choose the stationary and fully symmetric convolution-based covariance model proposed by Rodrigues & Diggle (2010), where negative, zero and positive non-separability can be achieved by varying a single parameter. Specifically, suppose the covariance function of a spatio-temporal process Z is given by

$$\mathrm{C}(\mathbf{h}, u) = \int_{\mathbb{R}^d} \int_{\mathbb{R}} k(\boldsymbol{\omega} - \mathbf{h}, \tau - u) k(\mathbf{h}, u) \mathrm{d}\boldsymbol{\omega} \mathrm{d}\tau, \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \qquad 13.$$

where k is a square-integrable kernel function. Rodrigues & Diggle (2010) suggested using

$$k(\mathbf{h}, u) = \rho(u) \exp\{-\rho(u)^\beta (\|\mathbf{h}\|/a)^2\}, \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \qquad 14.$$

where $\rho(u) = \exp(-|u|/b)$ is an exponential temporal correlation function with $b > 0$ as the temporal range parameter, $a > 0$ is the spatial range parameter, and $\beta < 2$ is the non-separability parameter. When $\beta = 0$, Equation 13 reduces to the separable case. Positive and negative values of β correspond to positive and negative non-separability, respectively, whose definitions are given in Section 3.1.3. The general form of the covariance function corresponding to the kernel k in Equation 14 is

$$\mathrm{C}_{\mathrm{RD}}(\|\mathbf{h}\|, u) \propto \int \frac{\rho(t-u)\rho(t)}{\rho(t-u)^\beta + \rho(t)^\beta} \exp\left\{ -\frac{\rho(t-u)^\beta \rho(t)^\beta}{\rho(t-u)^\beta + \rho(t)^\beta} \left(\frac{\|\mathbf{h}\|}{a} \right)^2 \right\} \mathrm{d}t, (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}. \qquad 15.$$

In the visuanimation, we choose values for the non-separability parameter, $\beta \in \{-1, 0, 1\}$, representing negative, zero and positive non-separability, respectively. In addition, we choose values for the spatial and temporal range parameters a and b corresponding to weak, medium and strong spatial and temporal dependences, measured by the respective effective ranges. The effective range is defined as the distance beyond which the correlation between observations is less than or equal to 0.05 (Irvine et al. 2007). In this section, we simulate at 15×15 locations in a unit square $[0,1]^2$ in space and 10 points in a unit interval $[0,1]$ in time, and select levels of the spatial effective range as $a^* \in \{0.3, 0.7, 1.2\}$, and the temporal effective range as $b^* \in \{0.3, 0.6, 0.9\}$. Fixing $\beta \in \{-1, 0, 1\}$, the corresponding values for a and b are computed by solving $\rho_{\mathrm{RD}}(0, b^*; a, b, \beta) = 0.05$ and $\rho_{\mathrm{RD}}(a^*, 0; a, b, \beta) = 0.05$, respectively, where $\rho_{\mathrm{RD}}(\|\mathbf{h}\|, u; a, b, \beta) = \mathrm{C}_{\mathrm{RD}}(\|\mathbf{h}\|, u)/\mathrm{C}_{\mathrm{RD}}(0,0)$ is the space-time correlation function, $\rho_{\mathrm{RD}}(0, u; a, b, \beta)$ and $\rho_{\mathrm{RD}}(\|\mathbf{h}\|, 0; a, b, \beta)$ are the purely temporal and purely spatial correlation functions, respectively. The complicated integral in Equation 15 is evaluated numerically.

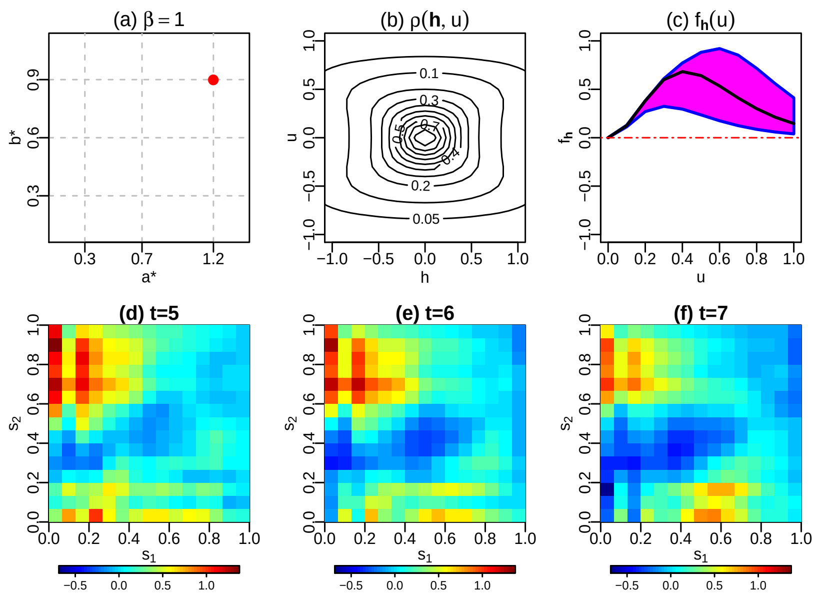

Figure 1

Illustration of the positive ($\beta > 0$) non-separability correlation structure. For each combination of parameters shown in panel (a), the contour plot of the correlation function is shown in (b), the functional boxplot of the true separability test functions (see Equation 6) is shown in (c), where each curve is associated with a specific pair of spatial locations, the central black curve represents the median, the magenta area is the 50% central region, and the band bounded by the outer curves is the maximum non-outlying envelope, and (d)-(f) show spatial maps at three consecutive time points from realizations of a spatio-temporal Gaussian process with zero mean, unit variance and the corresponding correlation function.

We display the visuanimations in Movie 1 in the supplementary materials, and here we only show one case in **Figure 1** for illustration. For each combination of parameters shown in panel (a), the contour plot of the correlation function is shown in (b), the functional boxplot of the true separability test functions (see Equation 6) defined by Huang & Sun

(2019) is shown in (c), where the degree of non-separability is indicated by the extent of departure from zero at all temporal lags, and (d)-(f) show spatial maps at three consecutive time points from realizations of a spatio-temporal Gaussian process with zero mean, unit variance and the corresponding correlation function. Movie 1 shows that the contour plots for all parameter combinations are symmetric, and the spatial correlation decays quickly if the effective range a^* is small, and decays slowly if a^* is big. When $\beta = 0$, the separability test functions are identically zero, and when $\beta \neq 0$, the test functions deviate from zero, or specifically, they are all negative when $\beta = -1$ and all positive when $\beta = 1$. The realizations show features that reflect the underlying correlation structures. When $\beta = 0$, i.e., in the separable case, the maps at the three consecutive time points do not interact, although in a few cases some spatial patterns are retained as time evolves, due to the strong temporal dependence. When $\beta \neq 0$, space-time interactions exist, i.e., the maps largely share spatial patterns at the three consecutive times, especially when the spatial or temporal dependence is strong.

In order to illustrate the fully symmetric and asymmetric covariance structures, we use a construction similar to that in Gneiting et al. (2007). Specifically, the correlation model is a linear combination of ρ_{RD} used above, and a Lagrangian correlation function ρ_{LGR},

$$\rho(\mathbf{h}, u) = (1-\lambda)\rho_{\mathrm{RD}}(\|\mathbf{h}\|, u) + \lambda\rho_{\mathrm{LGR}}(\mathbf{h}, u), \qquad 16.$$

where $\lambda \in [0, 1]$ is the asymmetry parameter, with $\lambda = 0$ indicating full symmetry, and

$$\rho_{\mathrm{LGR}}(\mathbf{h}, u) = \left(1 - \frac{1}{2\|\mathbf{v}\|}\|\mathbf{h} - \mathbf{v}u\|\right)_+^{3/2}, \qquad 17.$$

where $p_+ = \max(p, 0)$, $\mathbf{h} = (h_1, h_2)^\top$, and $\mathbf{v} = (v_1, v_2)^\top$ is the vector of constant velocities at the two directions h_1 and h_2, respectively. The Lagrangian correlation function ρ_{LGR} is valid since in Equation 12, we set $\mathbf{V} = \mathbf{v}$ as a constant vector, and use the Askey function (Askey 1973), $\mathrm{C_S}(\mathbf{h}) = (1 - \|\mathbf{h}\|/b)_+^{\nu}$, $b > 0$, which is positive definite on $\mathbb{R}^2$ if $\nu > 3/2$ (Zastavnyi & Trigub 2002); also, we set $b = 2\|\mathbf{v}\| > 0$.

For visuanimation, we fix the non-separability parameter $\beta = 1$, and choose different values for the asymmetry parameter, $\lambda \in \{0, 0.05, 0.1\}$. When $\lambda = 0$, the covariance function reduces to the fully symmetric case, and it is isotropic, i.e., it depends on the spatial and temporal lags only through their distances. When $\lambda \neq 0$, the covariance function is anisotropic, so we cannot deduce the parameter settings that correspond to specific configurations of effective ranges, as in the first illustration. However, since here λ is small, the spatial and temporal dependences are affected by the Lagrangian part only slightly. Hence, for $\lambda = 0.05$ or $\lambda = 0.1$, we use the same parameter settings that are derived from the case where $\lambda = 0$. We use the same configuration for simulation as above, and assume $v_1 = v_2 = 1$, i.e., the random field travels in the direction of 45°. The visuanimations can be found in Movie 2 in the supplementary materials, and only one case is shown here in **Figure 2** for illustration. Since the correlation function relies on the temporal lag u and both components, h_1 and h_2, of the spatial lag $\mathbf{h}$, we fix $h_2 = 0$, and draw the contours for different values of u and h_1. We see that when $\lambda = 0$, the contour plots in (b) are symmetric, and the full symmetry test functions in (c) are identically zero; when $\lambda > 0$, the contour plots are asymmetric, and the test functions deviate from zero, with the extent of asymmetry and departure from zero getting larger as λ increases. In addition, when $\lambda = 0$, the maps are the same as those in the case of $\beta = 1$ in Movie 1, since we use the

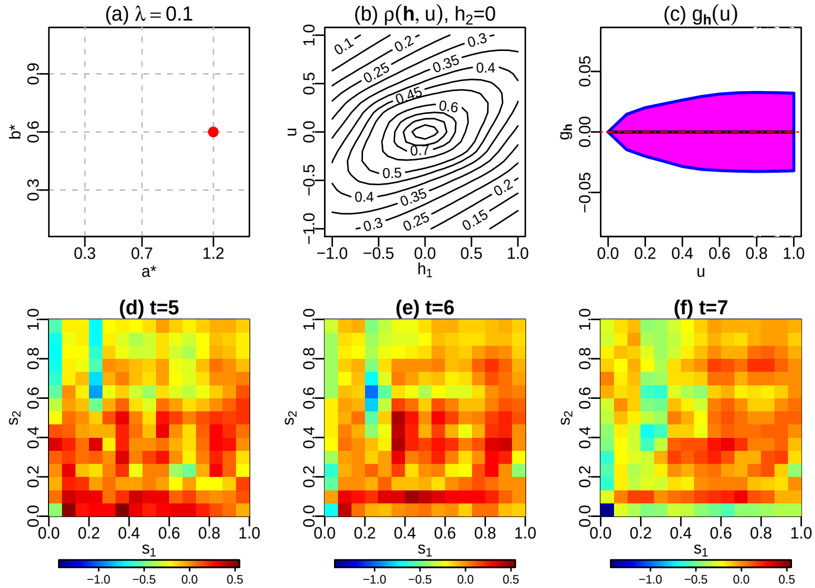

Figure 2

Illustration of the asymmetric ($0 < \lambda < 1$) correlation structure. For each combination of parameters shown in panel (a), the contour plot of the correlation function, with h_2 fixed at zero, is shown in (b), the functional boxplot of the full symmetry test functions (see Equation 6) is shown in (c), and (d)-(f) show spatial maps at three consecutive time points from realizations of a spatio-temporal Gaussian process with zero mean, unit variance and the corresponding correlation function.

same covariance model and the same seed in simulating the Gaussian random field. When $\lambda > 0$, a transport effect is present in the direction of 45°, and becomes more obvious as λ increases.

4. FITTING SPACE-TIME COVARIANCE MODELS

The estimation of covariance functions is important for applications of space-time random processes, such as space-time prediction, or kriging. Covariance functions can be estimated non-parametrically without assuming any covariance model. For example, a stationary covariance function can be estimated empirically by

$$\widehat{\mathrm{C}}(\mathbf{h}, u) = \frac{\sum_{\mathbf{s},\mathbf{s}+\mathbf{h}\in D} \sum_{i=1}^{T-u} \left\{ Z(\mathbf{s}, i) - \frac{1}{T-u} \sum_{j=1}^{T-u} Z(\mathbf{s}, j) \right\} \left\{ Z(\mathbf{s}+\mathbf{h}, i+u) - \frac{1}{T-u} \sum_{j=1}^{T-u} Z(\mathbf{s}+\mathbf{h}, j+u) \right\}}{N(\mathbf{h})(T-u)}, \quad 18.$$

for $u = 0, \ldots, T-1$, where T is the number of time points, $N(\mathbf{h})$ is the number of pairs of locations with spatial lag $\mathbf{h}$, and D is the spatial domain of the observations. More often, one selects from various parametric classes of covariance models and estimates the parameters therein by fitting the models to the data. This allows one to learn the properties of the covariance in the data, such as separability, full symmetry, and the rates of decay of spatial and temporal correlations, from the estimated parameters. Furthermore, stochastic weather generators (SWG, Wilks & Wilby 1999) based on fast statistical simulations with the estimated parameters can be used to reproduce the process, and spatio-temporal kriging can be performed to predict the process in space and time.

After choosing a model, the parameters can be estimated by using an ordinary, weighted or generalized least-squares approach based on empirical and fitted variograms. Under the typical Gaussian assumptions, the maximum likelihood (ML) approach is statistically most efficient for estimation. Suppose we have a vector of space-time observations $\mathbf{Z} = \{Z(\mathbf{s}_1, t_1), \ldots, Z(\mathbf{s}_N, t_N)\}^\top$ with the mean removed so that we focus on the estimation of the covariance. Assume that $\mathbf{Z} \sim \mathcal{N}_N(\mathbf{0}, \mathbf{\Sigma}(\theta))$, where θ is the vector of parameters in the covariance model, and $\mathbf{\Sigma}(\theta)_{ij} = \text{cov}\{Z(\mathbf{s}_i, t_i), Z(\mathbf{s}_j, t_j)\}$, $i, j = 1, \ldots, N$. The log-likelihood function is given by

$$l(\theta) \propto -\ln[\det\{\mathbf{\Sigma}(\theta)\}] - \mathbf{Z}^\top \mathbf{\Sigma}(\theta)^{-1}\mathbf{Z}. \qquad 19.$$

The ML estimator is then obtained by maximizing $l(\theta)$. However, exact ML estimation is computationally infeasible for large datasets, since it requires the evaluation of the determinant and inverse of an $N \times N$ matrix, which has a computational complexity of $\mathcal{O}(N^3)$.

This scalability issue has also been recognized in spatial statistics. To speed up the computation, one estimation technique is based on computationally tractable approximations of the covariance matrix. There are three common strategies in spatial statistics: low rank, sparse and spectral approximations. The low rank method approximates $\mathbf{\Sigma}$ with a matrix of smaller rank, m, reducing the computational complexity to $\mathcal{O}(nm^2)$, where n is the number of spatial locations. Basis function representations are often used for this purpose; see, e.g., Cressie & Johannesson (2008) and Katzfuss (2017). The sparse method introduces sparsity in the full covariance matrix through a compactly supported covariance representation, i.e., the covariance shrinks to zero whenever the spatial and/or temporal distances are sufficiently large. The covariance tapering method (Kaufman et al. 2008), for instance, belongs to this category. The Gaussian Markov random field approximation (Rue & Tjelmeland 2002), on the other hand, introduces sparsity in the inverse of the covariance matrix (also called the precision matrix). When the observations are located on a regular lattice, the spectral representation of the Gaussian process allows one to develop spectral constructions of the full covariance matrix. For example, Samo & Roberts (2015) proposed generalized spectral kernels that can approximate any bounded kernel with arbitrary precision.

Another strategy is to use a simplified version of the full likelihood, such as composite likelihood (CL) (Lindsay 1988); see Varin et al. (2011) for an overview. The CL is a product of a collection of component likelihoods. If $\{\mathcal{A}_1, \ldots, \mathcal{A}_K\}$ is a set of marginal or conditional events with associated likelihoods $\mathcal{L}_k(\theta) \propto f(\mathbf{z} \in \mathcal{A}_k; \theta)$, then a CL is the weighted product: $\mathcal{L}_\text{C}(\theta) = \prod_{k=1}^K \mathcal{L}_k(\theta)^{\omega_k}$, where ω_k's are non-negative weights to be chosen. In particular applications, unequal weights can be chosen to improve the efficiency. Because each individual component is a conditional or marginal density, the estimating equation obtained from the derivative of the CL is unbiased. However, because the components are multiplied, whether or not they are independent, the inference function has the properties of a likelihood from a misspecified model.

It is expected that the methods discussed above can be extended to the spatio-temporal case. However, additional challenges arise when analyzing spatio-temporal data, due to the high dimensionality with the addition of time, and the distinct, yet intricately involved, nature of space and time. To simplify the analysis, some authors either separately model the spatial and temporal dependences (Sahu et al. 2007), or apply a separable space-time covariance function (Genton 2007). However, these approaches ignore the crucial effect of space-time interaction. Paciorek et al. (2009) attempted to capture the space-time interaction of PM10- and PM2.5-particles by using monthly varying spatial surfaces, but assumed independence across spatial residual surfaces at each time point, which limited their ability to quantify the space-time interaction. Bai et al. (2012) developed an efficient CL approach for the joint analysis of spatio-temporal processes. They first constructed three sets of estimating functions from spatial, temporal and spatio-temporal cross-pairs, which result in over-identified estimating functions, and then formed a joint inference function in a spirit that is similar to the generalized method of moments (Hansen 1982). Bevilacqua et al. (2012) proposed two estimation methods using the CL approach. The first relies on the maximization of a weighted version of the CL function with cutoff weights based on the distance in space and in time. The second is based on the solution of a weighted CL equation, where the weights are obtained by minimizing an upper bound for the asymptotic variance of the estimator. Porcu et al. (2020) proposed spatio-temporal covariance models with dynamical compact support, meaning that the compact supports depend on the spatial and temporal lags. Their model brings sparsity in the covariance matrix, and the sparsity changes at each iteration of the maximization algorithm in the Gaussian likelihood computation.

Recently, software called `ExaGeoStat` (Abdulah et al. 2018) has been developed for Gaussian spatial processes. It allows for exact ML estimation with dense full covariance matrices, using high performance computations that employ the most advanced parallel architectures, combined with cutting edge dense linear algebra libraries. `ExaGeoStat` was also fine-tuned to work on the tile low-rank representation of the dense full covariance matrix. Development of `ExaGeoStat` for spatio-temporal settings, however, is ongoing.

The recent trend of using hierarchical models could also ease the computational burden, but in a completely different way; see Cressie & Wikle (2011) for more details. When there is strong evidence showing non-Gaussianity or non-stationarity, one can refer to the methods in, e.g., de Luna & Genton (2005), Wikle & Royle (2005) and Fonseca & Steel (2011b).

5. SPACE-TIME DATA ANALYSIS

In this section, we analyze the space-time covariance structure of a wind dataset, which was produced by Yip (2018), using the Weather Research and Forecasting (WRF) model. The original data consist of hourly wind speeds at a fine spatial resolution of approximately 5 km $\times$ 5 km covering the Arabian Peninsula, over the period 2009–2014. We select a subregion in the new mega-city, NEOM, of Saudi Arabia (NEOM Project 2017), bounded approximately by longitudes 35.49–35.67°E and latitudes 28.23–28.41°N. The NEOM project, initiated from the Saudi Vision 2030 (Alturki & Alsheikh 2016), envisages a very large, self-sustainable city with a substantial reliance on wind energy. For the time domain, we analyze the hourly winds in the week of August 20 to 26, 2014, characterized by one of the strongest wind regimes of the year (Chen et al. 2018). The final spatio-temporal dataset we consider consists of wind speeds at $n = 5 \times 5 = 25$ gridded spatial locations and $T = 24 \times 7 = 168$ time points, yielding a total of $nT = 4200$ data points.

Following common practice in the literature (e.g., Gneiting et al. 2007), we apply the square root transform to the wind speeds. This stabilizes the variance over both space and time, and makes the marginal distributions approximately normal. We then fit a time series model for each location, and remove the location-specific trend and seasonality. The residuals, also called velocity measures, at the 25 locations are denoted by

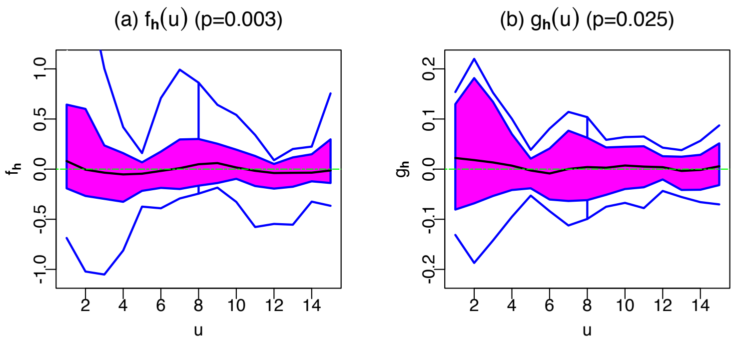

Figure 3

Functional boxplots for (a) the separability and (b) the full symmetry test functions defined in Equation 6 for the velocity measures and the p-values of the rank-based tests.

$\mathbf{Z} = (Z_{11}, Z_{21}, \ldots, Z_{n1}, Z_{12}, \ldots, Z_{n2}, \ldots, Z_{1T}, \ldots, Z_{nT})^\top$.

Next, we fit a spatio-temporal Gaussian process with zero mean and a space-time covariance function to the velocity measures $\mathbf{Z}$. Prior to choosing a covariance model, we use the testing procedure developed in Huang & Sun (2019) to assess the separability and full symmetry properties of the covariance function for $\mathbf{Z}$, calculated by Equation 18. **Figure 3** shows the functional boxplots for the two test functions, which clearly deviate from zero at all temporal lags, indicating non-separability and asymmetry. The significance of non-separability and asymmetry can be assessed using the rank-based tests, which yield p-values of 0.003 and 0.025, respectively, based on 1000 bootstrap replicates. These imply that the assumptions of separability and full symmetry are violated.

Therefore, we use a stationary but not necessarily fully symmetric covariance model similar to that used in Gneiting et al. (2007). The correlation model is a linear combination of the Gneiting (2002) model and the frozen Lagrangian model in Equation 17,

$$\rho(\mathbf{h}, u) = \frac{1-\lambda}{a|u|^{2\nu}+1} \exp\left\{-\frac{b\|\mathbf{h}\|}{(a|u|^{2\nu}+1)^{\beta/2}}\right\} + \lambda\left(1 - \frac{\|\mathbf{h}-\mathbf{v}u\|}{2\|\mathbf{v}\|}\right)_+^{3/2}, \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \qquad 20.$$

where $\lambda \in [0, 1]$ is the asymmetry parameter, $a > 0$ and $b > 0$ are temporal and spatial range parameters, respectively, $\nu > 0$ is the smoothness parameter, $\beta \in [0, 1]$ is the non-separability parameter, and $\mathbf{v}$ is the constant velocity vector. We also consider a joint spatio-temporal nugget effect, δ, and a sill parameter σ^2, so the full parametric covariance model is

$$\mathrm{C}^*(\mathbf{h}, u) = \sigma^2\rho(\mathbf{h}, u) + \delta I\{(\mathbf{h}, u) = (\mathbf{0}, 0)\}, \quad (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \qquad 21.$$

where I is the indicator function, and ρ is given by Equation 20.

Under Gaussianity, we use the ML method to estimate the parameters. Since the likelihood surface would be multimodal if we jointly estimate all the parameters, we fix the asymmetry parameter at four levels, $\lambda \in \{0, 0.05, 0.1, 0.15\}$. Covariance tapering is exploited to reduce the computational burden, with the empirical spatial and temporal effective ranges used to determine the range of the compact support for the taper functions. Since the empirical spatial correlations exceed 0.2 for all pairs of locations, we do not use a

Table 1 ML estimates for the covariance parameters in a symmetric model ($\lambda = 0$) and three asymmetric models ($\lambda = 0.05, 0.1, 0.15$). The standard errors based on the Hessian matrix for the best fitting case are shown in brackets below the estimates.

	$\widehat{\beta}$	$\widehat{a}$	$\widehat{b}$	$\widehat{\nu}$	$\widehat{v}_1$	$\widehat{v}_2$	$\widehat{\sigma}^2$	$\widehat{\delta}$	AIC
$\lambda = 0$	0.12	3.20	2.92	30.61	-	-	0.14	0.08	$-5,322$
$\lambda = 0.05$	0.20	0.85	0.97	22.91	15.02	-224.19	0.17	0.07	$-9,451$
$\lambda = 0.10$	0.14	0.69	1.80	34.07	21.07	-293.26	0.09	0.03	$-12,587$
	(0.01)	(0.04)	(0.34)	(8.27)	(3.04)	(25.30)	(0.01)	(0.002)	
$\lambda = 0.15$	0.18	0.85	1.72	27.48	17.88	-315.17	0.09	0.05	$-11,758$

spatial taper. The empirical temporal correlation reduces to 0.05 when the temporal lag is 15 hours, so we use a temporal taper, C_T, that belongs to the Wendland class (Wendland 1995) of compactly supported covariance functions, i.e., $C_T(u) = (1 - |u|/\theta_T)_+^4(1 + 4|u|/\theta_T), u \in \mathbb{R}$, where $\theta_T = 15$ hours. Then the tapered covariance matrix is obtained as $C(\mathbf{h}, u) = C^*(\mathbf{h}, u)C_T(u), (\mathbf{h}, u) \in \mathbb{R}^d \times \mathbb{R}$, where C^* is given by Equation 21. The marginal temporal correlations are zero whenever the temporal lag is greater than $\theta_T = 15$ hours. This results in a proportion of 83.5% zero values in the associated covariance matrix. Gaussian likelihood optimization is then performed exploiting algorithms for sparse matrices as implemented in the package `spam` (Furrer & Sain 2010) in R (R Core Team 2020). The estimated parameters are shown in **Table 1**. Based on the Akaike information criterion (AIC), the asymmetric model with $\lambda = 0.1$ gives the best fit. In addition, comparing $\hat{a}$ and $\hat{b}$ in the symmetric and asymmetric models, we see that the estimated spatial and temporal dependences are stronger in the asymmetric model.

We have chosen covariance tapering for illustration purposes. Other approximation methods are available for the likelihood computation. We did not perform prediction for this application due to its complexity and the limited length of this paper. R packages such as `gstat` (Pebesma 2004), `spacetime` (Pebesma 2012) and `CompRandFld` (Padoan & Bevilacqua 2015) can be used for inference and prediction, but only a limited number of classical covariance models, either separable or non-separable, have been implemented, and none of them includes asymmetric models.

SUMMARY POINTS

1. The positive definiteness condition for a function to be a covariance makes it non-trivial to construct covariance models, especially in the space-time setting.
2. Simplified structures, including stationarity, isotropy, separability, full symmetry and Taylor's hypothesis, can be imposed on the space-time covariance for the purposes of construction and computational efficiency.
3. Formal testing procedures can be used to validate these assumptions.
4. Separability does not allow for space-time interaction in the covariance, so there is a major demand for non-separable models, which can be constructed from Fourier transform of permissible spectral densities, mixtures of separable models, and partial differential equations representing physical laws.
5. The Lagrangian reference frame is useful for constructing asymmetric space-time covariance models.

6. Geodesic distance is the physically most natural metric for processes on spheres. The extension of covariance models using geodesic distance to spatio-temporal, multivariate, non-stationary, or non-Gaussian cases is not well understood.
7. The scalability issue can be addressed by approximating the covariance matrix using low rank, sparse and spectral techniques, or by using composite likelihood methods.

FUTURE ISSUES

1. Developing flexible space-time covariance models for multivariate, non-stationary, non-Gaussian, and spherical processes based on the geodesic distance.
2. Proposing tractable measures of non-separability and covariance models that can achieve weak to strong degrees of non-separability.
3. Developing easy-to-use software for inference from large spatio-temporal datasets with implementations of various covariance models, including non-separable, asymmetric, non-stationary and multivariate models, as well as on the sphere.
4. Proposing more efficient strategies and algorithms for fast inference and predictions with space-time covariance models for large datasets.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This publication is based on research supported by the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award No: OSR-2018-CRG7-3742.

LITERATURE CITED

Abdulah S, Ltaief H, Sun Y, Genton MG, Keyes DE. 2018. ExaGeoStat: A high performance unified software for geostatistics on manycore systems. *IEEE Trans. Parallel Distrib. Syst.* 29:2771–2784

Ailliot P, Baxevani A, Cuzol A, Monbet V, Raillard N. 2011. Space–time models for moving fields with an application to significant wave height fields. *Environmetrics* 22:354–369

Alegría A, Porcu E. 2017. The dimple problem related to space–time modeling under the Lagrangian framework. *J. Multivar. Anal.* 162:110–121

Alegría A, Porcu E, Furrer R, Mateu J. 2019. Covariance functions for multivariate Gaussian fields evolving temporally over planet earth. *Stoch. Environ. Res. Risk. Assess.* 33:1593–1608

Alturki FM, Alsheikh R. 2016. Saudi Vision 2030. `http://vision2030.gov.sa/en/node/87`. Accessed: 02-28-2020

Apanasovich TV, Genton MG. 2010. Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika* 97:15–30

Askey R. 1973. Radial characteristic functions. *Tech. rep.*, Wisconsin University, Madison Mathematics Research Center

Aston JA, Pigoli D, Tavakoli S. 2017. Tests for separability in nonparametric covariance operators of random surfaces. *Ann. Stat.* 45:1431–1461

Bai Y, Song PXK, Raghunathan T. 2012. Joint composite estimating functions in spatiotemporal models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 74:799–824

Bandyopadhyay S, Jentsch C, Subba Rao S. 2017. A spectral domain test for stationarity of spatio-temporal data. *J. Time Ser. Anal.* 38:326–351

Bandyopadhyay S, Rao SS. 2017. A test for stationarity for irregularly spaced spatial data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 79:95–123

Berg C, Porcu E. 2017. From Schoenberg coefficients to Schoenberg functions. *Constr. Approx.* 45:217–241

Bevilacqua M, Gaetan C, Mateu J, Porcu E. 2012. Estimating space and space-time covariance functions for large data sets: a weighted composite likelihood approach. *J. Am. Stat. Assoc.* 107:268–280

Bevilacqua M, Mateu J, Porcu E, Zhang H, Zini A. 2010. Weighted composite likelihood-based tests for space-time separability of covariance functions. *Stat. Comput.* 20:283–293

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

Bornn L, Shaddick G, Zidek JV. 2012. Modeling nonstationary processes through dimension expansion. *J. Am. Stat. Assoc.* 107:281–289

Bourotte M, Allard D, Porcu E. 2016. A flexible class of non-separable cross-covariance functions for multivariate space–time data. *Spat. Stat.* 18:125–146

Brown PE, Diggle PJ, Lord ME, Young PC. 2001. Space–time calibration of radar rainfall data. *J. R. Stat. Soc. Ser. C Appl. Stat.* 50:221–241

Brown PE, Roberts GO, Kåresen KF, Tonellato S. 2000. Blur-generated non-separable space–time models. *J. R. Stat. Soc. Series B Stat. Methodol.* 62:847–860

Cappello C, De Iaco S, Posa D. 2018. Testing the type of non-separability and some classes of space–time covariance function models. *Stoch. Environ. Res. Risk. Assess.* 32:17–35

Chaganty NR, Naik DN. 2002. Analysis of multivariate longitudinal data using quasi-least squares. *J. Stat. Plan. Inference* 103:421–436

Chen W, Castruccio S, Genton MG, Crippa P. 2018. Current and future estimates of wind energy potential over Saudi Arabia. *J. Geophys. Res. Atmos.* 123:6443–6459

Christakos G. 1991. On certain classes of spatiotemporal random fields with applications to space-time data processing. *IEEE Trans. Syst. Man. Cybernet.* 21:861–875

Christakos G. 2000. *Modern Spatiotemporal Geostatistics.* Oxford University Press

Christakos G. 2004. *Random Field Models in Earth Sciences.* San Diego, CA: Academic Press; 1992 (out of print). New edition, Mineola, NY: Dover Publ. Inc.

Christakos G. 2017. *Spatiotemporal Random Fields: Theory and Applications.* Elsevier

Christakos G, Hristopulos D. 1998. *Spatiotemporal Environmental Health Modelling.* Dordrecht, The Netherlands: Kluwer Acad.

Constantinou P, Kokoszka P, Reimherr M. 2017. Testing separability of space-time functional processes. *Biometrika* 104:425–437

Cox DR, Isham V. 1988. A simple spatial-temporal model of rainfall. *Proc. R. Soc. London, Ser. A* 415:317–328

Cressie N, Huang HC. 1999. Classes of nonseparable, spatio-temporal stationary covariance functions. *J. Am. Stat. Assoc.* 94:1330–1339

Cressie N, Johannesson G. 2008. Fixed rank kriging for very large spatial data sets. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 70:209–226

Cressie N, Wikle CK. 2011. *Statistics for Spatio-Temporal Data.* New York: Wiley

Crujeiras RM, Fernández-Casal R, González-Manteiga W. 2010. Nonparametric test for separability of spatio-temporal processes. *Environmetrics* 21:382–399

De Iaco S, Myers D, Posa D. 2011. On strict positive definiteness of product and product–sum covariance models. *J. Stat. Plan. Inference* 141:1132–1140

De Iaco S, Myers DE, Posa D. 2001. Space–time analysis using a general product–sum model. *Stat. Probab. Lett.* 52:21–28

De Iaco S, Myers DE, Posa D. 2002. Nonseparable space-time covariance models: some parametric families. *Math. Geol.* 34:23–42

De Iaco S, Palma M, Posa D. 2019. Choosing suitable linear coregionalization models for spatio-temporal data. *Stoch. Environ. Res. Risk. Assess.* 33:1419–1434

De Iaco S, Posa D. 2013. Positive and negative non-separability for space–time covariance models. *J. Stat. Plan. Inference* 143:378–391

de Luna X, Genton MG. 2005. Predictive spatio-temporal models for spatially sparse enviromental data. *Stat. Sin.* 15:547–568

Ezzat AA, Jun M, Ding Y. 2018. Spatio-temporal asymmetry of local wind fields and its impact on short-term wind forecasting. *IEEE Trans. Sustain. Energy* 9:1437–1447

Fonseca TC, Steel MF. 2011a. A general class of nonseparable space–time covariance models. *Environmetrics* 22:224–242

Fonseca TC, Steel MF. 2011b. Non-Gaussian spatiotemporal modelling through scale mixing. *Biometrika* 98:761–774

Fonseca TC, Steel MF. 2017. Measuring separability in spatio–temporal covariance functions. *Working paper*

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

Fuentes M. 2005. A formal test for nonstationarity of spatial stochastic processes. *J. Multivar. Anal.* 96:30–54

Fuentes M. 2006. Testing for separability of spatial–temporal covariance functions. *J. Stat. Plan. Inference* 136:447–466

Fuentes M, Chen L, Davis JM. 2008. A class of nonseparable and nonstationary spatial temporal covariance functions. *Environmetrics* 19:487–507

Fuentes M, Smith RL. 2001. A new class of nonstationary spatial models. *Tech. rep.*, North Carolina State University. Dept. of Statistics

Furrer R, Sain SR. 2010. spam: A sparse matrix R package with emphasis on MCMC methods for Gaussian Markov random fields. *J. Stat. Softw.* 36:1–25

Genton MG. 2007. Separable approximations of space-time covariance matrices. *Environmetrics* 18:681–695

Genton MG, Castruccio S, Crippa P, Dutta S, Huser R, Sun, Y, Vettori, S. 2015. Visuanimation in statistics. *Stat* 4:81–96

Genton MG, Kleiber W. 2015. Cross-covariance functions for multivariate geostatistics (with discussion). *Stat. Sci.* 30:147–163

Gneiting T. 2002. Nonseparable, stationary covariance functions for space–time data. *J. Am. Stat. Assoc.* 97:590–600

Gneiting T. 2013. Strictly and non-strictly positive definite functions on spheres. *Bernoulli* 19:1327–1349

Gneiting T, Genton MG, Guttorp P. 2007. Geostatistical space-time models, stationarity, separability, and full symmetry. In Finkenstaedt B, Held L, Isham V (eds.). *Statistics of Spatio-Temporal Systems,* Chapman & Hall/CRC Press, Monographs On Statistics and Applied Probability 151–175

Gneiting T, Kleiber W, Schlather M. 2010. Matérn cross-covariance functions for multivariate random fields. *J. Am. Stat. Assoc.* 105:1167–1177

Gneiting T, Schlather M. 2002. Space-time covariance models. In *Encyclopedia of Environmetrics*, eds. AH El-Shaarawi, WW Piegorsch, vol. 4. Chichester: Wiley, 2041–2045

Guinness J, Fuentes M. 2016. Isotropic covariance functions on spheres: Some properties and modeling considerations. *J. Multivar. Anal.* 143:143–152

Guo JH, Billard L. 1998. Some inference results for causal autoregressive processes on a plane. *J.*

Time Ser. Anal. 19:681–691
Gupta VK, Waymire E. 1987. On Taylor's hypothesis and dissipation in rainfall. *J. Geophys. Res. Atmos.* 92:9657–9660
Hansen LP. 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50(4):1029–1054
Heine V. 1955. Models for two-dimensional stationary stochastic processes. *Biometrika* 42:170–178
Higdon D, Swall J, Kern J. 1999. Non-stationary spatial modeling. In *Bayesian Statistics*, vol. 6, eds. J Bernardo, J Berger, A Dawid, A Smith. Oxford University Press, 761–768
Hitczenko M, Stein ML. 2012. Some theory for anisotropic processes on the sphere. *Stat. Method.* 9:211–227
Horrell MT, Stein ML. 2017. Half-spectral space–time covariance models. *Spat. Stat.* 19:90–100
Huang H, Sun Y. 2019. Visualization and assessment of spatio-temporal covariance properties. *Spat. Stat.* 34:100272
Huizenga HM, De Munck JC, Waldorp LJ, Grasman RP. 2002. Spatiotemporal eeg/meg source analysis based on a parametric noise covariance model. *IEEE Trans. Biomed. Eng.* 49:533–539
Inoue T, Sasaki T, Washio T. 2012. Spatio-temporal kriging of solar radiation incorporating direction and speed of cloud movement, In *The 26th Annual Conference of the Japanese Society for Articial Intelligence*, vol. 26
Ip RH, Li WK. 2016. Matérn cross-covariance functions for bivariate spatio-temporal random fields. *Spat. Stat.* 17:22–37
Ip RH, Li WK. 2017a. A class of valid Matérn cross-covariance functions for multivariate spatio-temporal random fields. *Stat. Probab. Lett.* 130:115–119
Ip RH, Li WK. 2017b. On some Matérn covariance functions for spatio-temporal random fields. *Stat. Sin.* 27:805–822
Ip RH, Li WK. 2015. Time varying spatio-temporal covariance models. *Spat. Stat.* 14:269–285
Irvine KM, Gitelman AI, Hoeting JA. 2007. Spatial designs and properties of spatial correlation: effects on covariance estimation. *J. Agric. Biol. Environ. Statist.* 12:450–469
Jeong J, Jun M. 2015. A class of Matérn-like covariance functions for smooth processes on a sphere. *Spat. Stat.* 11:1–18
Jeong J, Jun M, Genton MG. 2017. Spherical process models for global spatial statistics. *Stat. Sci.* 32:501–513
Jones RH, Zhang Y. 1997. Models for continuous stationary space-time processes. In *Modelling Longitudinal and Spatially Correlated Data*. New York: Springer, 289–298
Jun M, Genton MG. 2012. A test for stationarity of spatio-temporal random fields on planar and spherical domains. *Stat. Sin.* 22:1737–1764
Jun M, Stein ML. 2007. An approach to producing space–time covariance functions on spheres. *Technometrics* 49:468–479
Jun M, Stein ML. 2008. Nonstationary covariance models for global data. *Ann. Appl. Stat.* 2:1271–1289
Katzfuss M. 2017. A multi-resolution approximation for massive spatial datasets. *J. Am. Stat. Assoc.* 112:201–214
Kaufman CG, Schervish MJ, Nychka DW. 2008. Covariance tapering for likelihood-based estimation in large spatial data sets. *J. Am. Stat. Assoc.* 103:1545–1555
Kolovos A, Christakos G, Hristopulos DT, Serre ML. 2004. Methods for generating non-separable spatiotemporal covariance models with potential environmental applications. *Adv. Water Resour.* 27:815–830
Kyriakidis PC, Journel AG. 1999. Geostatistical space–time models: a review. *Math. Geol.* 31:651–684
Li B, Genton MG, Sherman M. 2007. A nonparametric assessment of properties of space–time covariance functions. *J. Am. Stat. Assoc.* 102:736–744
Li B, Genton MG, Sherman M. 2008a. On the asymptotic joint distribution of sample space: Time

covariance estimators. *Bernoulli* 14:228–248
Li B, Genton MG, Sherman M. 2008b. Testing the covariance structure of multivariate random fields. *Biometrika* 95:813–829
Li B, Murthi A, Bowman KP, North GR, Genton MG, Sherman M. 2009. Statistical tests of Taylor's hypothesis: an application to precipitation fields. *J. Hydrometeorol.* 10:254–265
Lindsay BG. 1988. Composite likelihood methods. *Contemp. Math.* 80:221–239
Liu C, Ray S, Hooker G. 2017. Functional principal component analysis of spatially correlated data. *Stat. Comput.* 27:1639–1654
Lonij VP, Brooks AE, Cronin AD, Leuthold M, Koch K. 2013. Intra-hour forecasts of solar power production using measurements from a network of irradiance sensors. *Sol. Energy* 97:58–66
Lu N, Zimmerman DL. 2005a. The likelihood ratio test for a separable covariance matrix. *Stat. Probab. Lett.* 73:449–457
Lu N, Zimmerman DL. 2005b. Testing for directional symmetry in spatial dependence using the periodogram. *J. Stat. Plan. Inference* 129:369–385
Ma C. 2002. Spatio-temporal covariance functions generated by mixtures. *Math. Geol.* 34:965–975
Ma C. 2003a. Families of spatio-temporal stationary covariance models. *J. Stat. Plan. Inference* 116:489–501
Ma C. 2003b. Spatio-temporal stationary covariance models. *J. Multivar. Anal.* 86:97–107
Ma C. 2005. Linear combinations of space-time covariance functions and variograms. *IEEE Trans. Signal Process.* 53:857–864
Matérn B. 1986. *Spatial Variation.* 2nd edition. New York: Springer
May DR, Julien PY. 1998. Eulerian and Lagrangian correlation structures of convective rainstorms. *Water Resour. Res.* 34:2671–2683
Mitchell M, Gumpertz M. 2003. Spatial variability inside a free-air CO2 system. *J. Agric. Biol. Environ. Statist.* 8:310–327
Mitchell MW, Genton MG, Gumpertz ML. 2005. Testing for separability of space–time covariances. *Environmetrics* 16:819–831
Mitchell MW, Genton MG, Gumpertz ML. 2006. A likelihood ratio test for separability of covariances. *J. Multivar. Anal.* 97:1025–1043
Montero JM, Fernández-Avilés G, Mateu J. 2015. *Spatial and Spatio-Temporal Geostatistical Modeling and Kriging*, New York: Wiley
NEOM Project. 2017. `https://www.neom.com/`. Accessed: 02-28-2020
North GR, Wang J, Genton MG. 2011. Correlation models for temperature fields. *J. Clim.* 24:5850–5862
Nychka D, Wikle C, Royle JA. 2002. Multiresolution models for nonstationary spatial covariance functions. *Stat. Model.* 2:315–331
Paciorek CJ, Yanosky JD, Puett RC, Laden F, Suh HH. 2009. Practical large-scale spatio-temporal modeling of particulate matter concentrations. *Ann. Appl. Stat.* 3:370–397
Padoan SA, Bevilacqua M. 2015. Analysis of random fields using CompRandFld. *J. Stat. Softw.* 63(9):1–27
Park MS, Fuentes M. 2008. Testing lack of symmetry in spatial–temporal processes. *J Stat. Plan. Inference* 138:2847–2866
Pebesma E. 2012. spacetime: Spatio-temporal data in R. *J. Stat. Softw.* 51:1–30
Pebesma EJ. 2004. Multivariable geostatistics in S: the gstat package. *Comput. Geosci.* 30:683–691
Pintore A, Holmes C. 2007. Spatially adaptive non-stationary covariance functions via spatially adaptive spectra. *Tech. rep.*, Oxford University, Statistics Department
Porcu E, Alegria A, Furrer R. 2018. Modeling temporally evolving and spatially globally dependent data. *Int. Stat. Rev.* 86:344–377
Porcu E, Bevilacqua M, Genton MG. 2016. Spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere. *J. Am. Stat. Assoc.* 111:888–898
Porcu E, Bevilacqua M, Genton MG. 2020. Nonseparable space-time covariance functions with

dynamical compact supports. *Stat. Sin.* 30:719–739

Porcu E, Gregori P, Mateu J. 2006. Nonseparable stationary anisotropic space–time covariance functions. *Stoch. Environ. Res. Risk. Assess.* 21:113–122

Porcu E, Mateu J, Bevilacqua M. 2007. Covariance functions that are stationary or nonstationary in space and stationary in time. *Stat. Neerl.* 61:358–382

Porcu E, Mateu J, Saura F. 2008. New classes of covariance and spectral density functions for spatio-temporal modelling. *Stoch. Environ. Res. Risk. Assess.* 22:65–79

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna: R Found. Stat. Comput. `http://www.R-project.org`

Rodrigues A, Diggle PJ. 2010. A class of convolution-based models for spatio-temporal processes with non-separable covariance structure. *Scand. J. Stat.* 37:553–567

Roy A, Khattree R. 2003. Tests for mean and covariance structures relevant in repeated measures based discriminant analysis. *J. Appl. Stat. Sci.* 12:91–104

Roy A, Khattree R. 2005a. On implementation of a test for Kronecker product covariance structure for multivariate repeated measures data. *Stat. Method.* 2:297–306

Roy A, Khattree R. 2005b. Testing the hypothesis of a Kronecker product covariance matrix in multivariate repeated measures data, In *Proceedings of the 30th Annual SAS Users Group International Conference, SUGI*

Roy A, Leiva R. 2008. Likelihood ratio tests for triply multivariate data with structured correlation on spatial repeated measurements. *Stat. Probab. Lett.* 78:1971–1980

Rue H, Tjelmeland H. 2002. Fitting Gaussian Markov random fields to Gaussian fields. *Scand. J. Stat.* 29:31–49

Sahu SK, Gelfand AE, Holland DM. 2007. High-resolution space–time ozone modeling for assessing trends. *J. Am. Stat. Assoc.* 102:1221–1234

Salvaña MLO, Genton MG. 2020. Nonstationary cross-covariance functions for multivariate spatio-temporal random fields. *Spat. Stat.* 37:100411

Samo YLK, Roberts S. 2015. Generalized spectral kernels. *arXiv preprint arXiv:1506.02236*

Sampson PD, Guttorp P. 1992. Nonparametric estimation of nonstationary spatial covariance structure. *J. Am. Stat. Assoc.* 87:108–119

Scaccia L, Martin R. 2002. Testing for simplification in spatial models, In *COMPSTAT*. New York: Springer

Scaccia L, Martin R. 2005. Testing axial symmetry and separability of lattice processes. *J. Stat. Plan. Inference* 131:19–39

Schlather M. 2010. Some covariance models based on normal scale mixtures. *Bernoulli* 16:780–797

Shand L, Li B. 2017. Modeling nonstationarity in space and time. *Biometrics* 73:759–768

Shao X, Li B. 2009. A tuning parameter free test for properties of space–time covariance functions. *J. Stat. Plan. Inference* 139:4031–4038

Shinozaki K, Yamakawa N, Sasaki T, Inoue T. 2016. Areal solar irradiance estimated by sparsely distributed observations of solar radiation. *IEEE Trans. Power Syst.* 31:35–42

Simpson SL. 2010. An adjusted likelihood ratio test for separability in unbalanced multivariate repeated measures data. *Stat. Method.* 7:511–519

Stein ML. 2005a. Space–time covariance functions. *J. Am. Stat. Assoc.* 100:310–321

Stein ML. 2005b. Statistical methods for regular monitoring data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 67:667–687

Sun Y, Genton MG. 2011. Functional boxplots. *J. Comput. Gr. Stat.* 20:316–334

Taylor GI. 1938. The spectrum of turbulence. *Proc. R. Soc. London, Ser. A* 164:476–490

Varin C, Reid N, Firth D. 2011. An overview of composite likelihood methods. *Stat. Sin.* 21:5–42

Weller ZD, Hoeting JA. 2016. A review of nonparametric hypothesis tests of isotropy properties in spatial data. *Stat. Sci.* 31:305–324

Wendland H. 1995. Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Adv. Comput. Math.* 4:389–396

Wikle CK, Royle JA. 2005. Dynamic design of ecological monitoring networks for non-Gaussian spatio-temporal data. *Environmetrics* 16:507–522

Wikle CK, Zammit-Mangion A, Cressie N. 2019. *Spatio-Temporal Statistics with R*. CRC Press

Wilks DS, Wilby RL. 1999. The weather generation game: a review of stochastic weather models. *Prog. Phys. Geogr.* 23:329–357

Xu H, Gardoni P. 2018. Improved latent space approach for modelling non-stationary spatial–temporal random fields. *Spat. Stat.* 23:160–181

Yaglom AM. 1987. *Correlation Theory of Stationary and Related Random Functions*, New York: Springer

Yip CMA. 2018. *Statistical Characteristics and Mapping of Near-Surface and Elevated Wind Resources in The Middle East*. Ph.D. thesis, King Abdullah University of Science and Technology

Zastavnyi VP, Trigub RM. 2002. Positive-definite splines of special form. *Sbornik Math.* 193:1771–1800