Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: http://www.tandfonline.com/loi/uasa20

Spatio-Temporal Covariance and Cross-Covariance Functions of the Great Circle Distance on a Sphere

Emilio Porcu, Moreno Bevilacqua & Marc G. Genton

To cite this article: Emilio Porcu, Moreno Bevilacqua & Marc G. Genton (2016) Spatio-Temporal Covariance and Cross-Covariance Functions of the Great Circle Distance on a Sphere, Journal of the American Statistical Association, 111:514, 888-898, DOI: 10.1080/01621459.2015.1072541

To link to this article: http://dx.doi.org/10.1080/01621459.2015.1072541

View supplementary material

Accepted author version posted online: 30 Jul 2015.
Published online: 18 Aug 2016.

Submit your article to this journal

Article views: 103

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=uasa20

Download by: [King Abdullah University of Science & Technology KAUST] **Date:** 18 August 2016, At: 15:36

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2016, VOL. 111, NO. 514, 888–898, Theory and Methods
http://dx.doi.org/10.1080/01621459.2015.1072541

Taylor & Francis
Taylor & Francis Group

Spatio-Temporal Covariance and Cross-Covariance Functions of the Great Circle Distance on a Sphere

Emilio Porcu, Moreno Bevilacqua, and Marc G. Genton

ABSTRACT

In this article, we propose stationary covariance functions for processes that evolve temporally over a sphere, as well as cross-covariance functions for multivariate random fields defined over a sphere. For such processes, the great circle distance is the natural metric that should be used to describe spatial dependence. Given the mathematical difficulties for the construction of covariance functions for processes defined over spheres cross time, approximations of the state of nature have been proposed in the literature by using the Euclidean (based on map projections) and the chordal distances. We present several methods of construction based on the great circle distance and provide closed-form expressions for both spatio-temporal and multivariate cases. A simulation study assesses the discrepancy between the great circle distance, chordal distance, and Euclidean distance based on a map projection both in terms of estimation and prediction in a space-time and a bivariate spatial setting, where the space is in this case the Earth. We revisit the analysis of Total Ozone Mapping Spectrometer (TOMS) data and investigate differences in terms of estimation and prediction between the aforementioned distance-based approaches. Both simulation and real data highlight sensible differences in terms of estimation of the spatial scale parameter. As far as prediction is concerned, the differences can be appreciated only when the interpoint distances are large, as demonstrated by an illustrative example. Supplementary materials for this article are available online.

ARTICLE HISTORY
Received July 2013
Revised May 2015

KEYWORDS
Chordal distance; Cokriging; Global data; Great circle distance; Multivariate; Spatio-temporal statistics; Sphere.

1. Introduction

Researchers in the environmental, geophysical, and agricultural sciences have become increasingly interested in global data covering a large portion of the Earth and in representing realizations of stochastic processes evolving temporally over a sphere. For such processes, Euclidean distance is not a valid metric for the description of spatial dependence, since it does not take into account the curvature of the Earth. In a recent tour de force, Gneiting (2013) gave an impressive overview of methods that can be used to generate spatial covariance functions depending on the great circle distance. He also described a collection of interesting and challenging open problems. This article addresses two of these problems and proposes solutions that are investigated in detail. Specifically, we propose spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere.

First, in his open Problem 16, Gneiting (2013) noted that "*Frequently, the temporal development of a process observed on a sphere is also of interest, so that the process needs to be modeled on the sphere cross time. Nevertheless, the literature on the corresponding correlation structures is sparse, with the work of Jun and Stein (2007) being a notable exception.*" Jun and Stein (2007) offered nonseparable spatio-temporal covariance functions whose spatial component depends on the chordal distance. It is a notable exception to a large literature entirely relying on the Euclidean distance based on map projections. Jun and Stein (2008) extended their approach to produce nonstationary covariances for data on a sphere, using once again the chordal distance.

Constructive criticisms in Gneiting (2013) suggest that the chordal distance approach has limited flexibility. For example, it does not allow negative correlations lower than -0.21 over space as can be seen from the structure of the integral representation of isotropic functions (Schoenberg 1938). Furthermore, the chordal distance is locally linear, and, according to Gneiting (2013), "*is counter to spherical geometry for larger values of the great circle distance, and thus may result in physically unrealistic distortions.*" For instance, North, Wang, and Genton (2011) developed spatio-temporal covariance models on a sphere for temperature fields based on simple yet realistic energy-balance climate models, but their construction is also based on chordal distances.

The great circle or orthodromic distance is the shortest distance between any two points on a sphere measured along a path on its surface. Because the geometry of the sphere is different from ordinary Euclidean geometry, the equations for distance take on a different form. The distance between two points in Euclidean space is the length of a straight line from one point to the other. On a sphere, however, there are no straight lines. In non-Euclidean geometry, straight lines are replaced by geodesics. On the sphere, geodesics are the great circles, that is, they are circles on the sphere whose centers coincide with the center of the sphere. Thus, the great circle distance is the most natural metric to account for phenomena evolving temporally over a sphere; see Banerjee (2005) for an overview in

CONTACT Emilio Porcu emilio.porcu@usm.cl Department of Mathematics, University Federico Santa Maria, 2360102 Valparaiso, Chile.
Color versions of one or more of the figures in the article can be found online at *www.tandfonline.com/r/JASA*.
Supplementary materials for this article are available online. Please go to *www.tandfonline.com/r/JASA*.
© 2016 American Statistical Association

spatial statistics. It is therefore imperative to propose new models of nonseparable covariance functions whose spatial argument depends on the great circle distance.

In the first part of the article, we address this problem from two different perspectives. With the first approach, that we call *adapted*, we show that a celebrated class of space-time covariance functions, called the *Gneiting class* (Gneiting 2002), and where the spatial component is based on Euclidean distance, can be adapted to covariances for processes on the sphere cross time, just by replacing the Euclidean distance with the great circle distance. Such covariances can also be adapted, by Yadrenko's (1983) principle, to chordal distance in the spatial component. In the same spirit, we show the validity of a new class, that we call *modified Gneiting class*, obtained with the same plugging technique as in Gneiting (2002), but where the temporal component does not rescale the spatial distance, which overcomes the dimple problem (Kent, Mohammadzadeh, and Mosammam 2011).

With the second approach, we show a constructive procedure for building space-time covariances that are valid on the sphere cross time but for which the Euclidean or the chordal distances cannot be used. Such an approach is more realistic from the point of view of the physical and mathematical construction of a Gaussian process on spheres cross time but will be shown to present some issues that might make the adaptive approach preferable in some cases.

A comparative study based on simulated data evolving over a sphere representing the Earth then shows how the chordal distance and the Euclidean distance based on a map projection (a common practice among geographic information system (GIS) users) cause distortions in the estimation of the spatial scale when a large portion of the Earth is considered. However, the variance and temporal scale are basically unaffected by the use of great circle or chordal distances. In the same spirit, we then address the problem of prediction over the sphere (see the online supplement, OS throughout) using cross-validation techniques, and we show that great circle and chordal distances perform equally well in terms of prediction. This is a fact that might be expected, since the accuracy of optimal linear prediction depends on the local properties of covariance functions (Stein 1999). We also use the proposed spatio-temporal models to analyze Level-3 Total Ozone Mapping Spectrometer (TOMS) data, which include the daily total column ozone levels, as studied with chordal distances and only purely spatial models by Jun and Stein (2008).

Environmental data are typically characterized by multivariate measurements at each site of the spatial domain; see, for example, Zhang (2007), Apanasovich and Genton (2010), Furrer and Genton (2011), Kleiber and Genton (2013), and the recent review by Genton and Kleiber (2015). In the case of global data over a large portion of the Earth, the arguments offered for justifying the use of the great circle distance obviously remain valid for the case of multivariate geostatistics. There is a very sparse literature on cross-covariance functions associated with vector-valued random fields defined over the sphere, with the work of Jun (2011) being an exception, albeit with the chordal distance again and not the great circle. Our article addresses this problem and proposes closed-form cross-covariance functions of the great circle distance. In particular, we propose multivariate Matérn mappings (Gneiting, Kleiber, and Schlather 2010; Apanasovich, Genton, and Sun 2012) that have been proposed earlier in the framework of Euclidean distance only. We also present a multivariate mapping of the Cauchy type that is new even for the Euclidean distance, as well as other examples.

With the same rationale as used for the univariate space-time case, we then perform a simulation study of a bivariate Gaussian field with a matrix-valued covariance of the Matérn type, and we show that even in this case, the choice of distance (great circle vs. chordal vs. Euclidean based on map projection) is a critical issue to estimation performance of the spatial scale.

In the sequel, we denote by $\mathbb{S}_R^d \subseteq \mathbb{R}^{d+1}$ the d-dimensional sphere with radius R, that is, $\mathbb{S}_R^d = \{\mathbf{x} \in \mathbb{R}^{d+1} : \|\mathbf{x}\| = R\}$. Throughout, we use the abuse of notation $\mathbb{S}^d$ for the unit sphere, and in what follows we refer to this case, whenever no confusion can arise. Let $\{\mathbf{Z}(\mathbf{x}, t) = \{Z_1(\mathbf{x}, t), \ldots, Z_p(\mathbf{x}, t)\}^{\mathrm{T}}, (\mathbf{x}, t) \in \mathbb{S}^d \times \mathbb{R}\}$ be a Gaussian p-dimensional vector-valued process defined on the sphere cross time. We denote by $\Psi_{d,T}^p$ the class of continuous square matrix-valued mappings:

$$\mathbf{C}(\cdot, \cdot) = \left[C_{ij}(\cdot, \cdot)\right]_{i,j=1}^p, \tag{1}$$

with $C_{ij} : [0, \pi] \times \mathbb{R} \to \mathbb{R}$, such that, for the Gaussian process $\mathbf{Z}$ defined above,

$$\operatorname{cov}\left\{\mathbf{Z}(\mathbf{x}, t), \mathbf{Z}(\mathbf{y}, t')\right\} = \mathbf{C}(\theta, t - t'), \qquad \theta \in [0, \pi], \quad t, t' \in \mathbb{R}, \tag{2}$$

where $\theta := \theta(\mathbf{x}, \mathbf{y}) = \arccos\left(\langle \mathbf{x}, \mathbf{y} \rangle\right)$ is the great circle distance between $\mathbf{x}$ and $\mathbf{y}$ on $\mathbb{S}^d$.

Such a description is parenthetical to the remark that $\mathbf{C}$ must be positive definite, that is, for any finite collection of real-valued vectors $\{\boldsymbol{a}_k\}_{k=1}^N$ and points $\{(\mathbf{x}_k, t_k) \in \mathbb{S}^d \times \mathbb{R}\}_{k=1}^N$, we have

$$\sum_{k,l=1}^{N} \sum_{i,j=1}^{p} a_{ki} a_{lj} C_{ij}\left\{\theta(\mathbf{x}_k, \mathbf{x}_l), |t_k - t_l|\right\} \geq 0.$$

For a univariate (scalar-valued) spatio-temporal process, we use the abuse of notation $\Psi_{d,T}$ for $\Psi_{d,T}^1$. For merely spatial processes observed over the sphere $\mathbb{S}^d$, we have $\operatorname{cov}\{\mathbf{Z}(\mathbf{x}), \mathbf{Z}(\mathbf{y})\} = \mathbf{C}(\theta)$, and we use the notation Ψ_d^p to describe this case. Obviously, the inclusion relation $\Psi_d^p \supset \Psi_{d,T}^p$ is strict, and this comes from the arguments in Corollary 1(b) in Gneiting (2013). The same arguments show that the following inclusion relation is strict:

$$\Psi_{1,T}^p \supset \Psi_{2,T}^p \supset \cdots \supset \bigcap_{d \geq 1} \Psi_{d,T}^p =: \Psi_{\infty,T}^p,$$

where $\Psi_{\infty,T}^p$ defines space-time matrix-valued covariances being valid on any d-dimensional sphere (actually a Hilbert sphere). Analogously, we use $\Psi_{d,T}$ and $\Psi_{\infty,T}$ for $\Psi_{d,T}^1$ and $\Psi_{\infty,T}^1$, respectively. Finally, we call $\Psi_d := \Psi_d^1$ the class of positive definite functions on the sphere $\mathbb{S}^d$, and $\Psi_\infty = \cap_{d \geq 1} \Psi_d$. Two more ingredients are necessary for the exposition of the subsequent results. A completely monotone function φ is a mapping from $[0, \infty)$ into $\mathbb{R}_+$ that is infinitely differentiable on $(0, \infty)$ and such that the derivatives change sign pointwise, that is, $(-1)^k \varphi^{(k)}(t) \geq 0$, for all $k \in \mathbb{N}, t > 0$. A function $f : [0, \infty) \to [0, \infty)$ is called a *Bernstein function* if it is infinitely differentiable on $(0, \infty)$ and its first derivative is completely monotone. For many interesting facts about these two classes of functions,

the reader is referred to Porcu and Schilling (2011) and the references therein.

The plan of this article is as follows. Section 2 deals with the class of univariate spatio-temporal covariance functions, $\Psi_{d,T}$. Section 3 deals with the class of cross-covariance functions of the great circle distance, Ψ_d^p $(p > 1)$. We then provide, in Section 4, a simulation study to emphasize the potential distortions in using the chordal distance or the Euclidean distance based on a map projection rather than the great circle distance in a space-time and a bivariate spatial setting. In Section 5, we revisit the TOMS data analysis proposed in Jun and Stein (2008). Proofs and technicalities are presented in the Appendix, reported in the online supplement.

2. Spatio-Temporal Covariance Functions of the Great Circle Distance: The Class $\Psi_{d,T}$

2.1 A Gneiting Class on the Sphere Cross Time: Adaptive Approach from the Euclidean Case

For a given space-time covariance whose spatial component depends on Euclidean distance, Yadrenko's principle (Yadrenko 1983) basically implies that we can readapt such a covariance to the sphere cross time, by replacing the Euclidean with the chordal distance. This fact prompts a natural question: can we replace the Euclidean distance argument with the great circle distance, while preserving positive definiteness? We call this approach *spatially adaptive*. We focus throughout on the Gneiting class (Gneiting 2002), which seems to be one of the most used in geostatistical applications; see also Gneiting, Genton, and Guttorp (2007) and references therein. For a Gaussian process Z defined on $\mathbb{R}^d \times \mathbb{R}$, we have that for two points $(\mathbf{x}, t)$ and $(\mathbf{y}, t')$ such that $\mathbf{y} - \mathbf{x} = \mathbf{h}$ (with $\|\cdot\|$ denoting the Euclidean distance) and $t - t' = u$, the Gneiting class of covariances is

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

where φ is completely monotone on the positive real line such that $\varphi(0) = 1$, ψ is a positive-valued Bernstein function, and σ^2 is a variance parameter. We remark that Gneiting (2002) presented a generalized version over the product space $\mathbb{R}^d \times \mathbb{R}^l$ and that, in the earlier literature, the spatial and temporal arguments have been inverted, but in the subsequent presentation, we prefer to work with such a parameterization. Zastavnyi and Porcu (2011) showed necessary and sufficient conditions for the positive definiteness of the Gneiting's class. The results below complete the picture of this class in the following way: we suppose that the spatial argument $\|\mathbf{h}\|$ is replaced with the great circle distance by restricting the function ψ in Equation (3) to the interval $[0, \pi]$. The proof and some technical lemmas are relegated to Appendix A1 (OS).

Theorem 1. Let $\theta : \mathbb{S}^d \times \mathbb{S}^d \to [0, \pi]$ be the great circle distance. Let $\varphi : [0, \infty) \to \mathbb{R}_+$ be a completely monotone function on the positive real line, with $\varphi(0) = 1$, and let ψ be a positive-valued Bernstein function. Denote by $\psi_{[0,\pi]}$ the restriction of ψ to the interval $[0, \pi]$. Then, the function

$$C(\theta, u) := \frac{\sigma^2}{\psi_{[0,\pi]}(\theta)^{1/2}} \varphi\left\{\frac{u^2}{\psi_{[0,\pi]}(\theta)}\right\}, \quad (\theta, u) \in [0, \pi] \times \mathbb{R}, \tag{4}$$

belongs to the class $\Psi_{\infty,T}$.

Theorem 1 offers a positive answer: we can work with a very general class of space-time covariances even on the sphere, just by readapting the spatial argument through the great circle distance. Following Gneiting (2002), we just need to plug into Equation (4) a valid choice of the functions φ and ψ. This can be easily done, for instance, from Tables S1 and S2 of the OS, where some cases are taken from Gneiting (2013) and Porcu and Schilling (2011), respectively. In Table S1, it may be surprising that the parameter ν in the Matérn function (second entry in Table S1 of the OS) is restricted to the interval $(0, 1/2]$, but it is carefully explained in Gneting (2013), on the basis of the arguments in Miller and Samko (2001), that such a family is not completely monotone for $\nu > 1/2$. Table S2 of the OS reports some possible choices for the function ψ (to be composed with the great circle distance). Some caution is needed when the value at zero is equal to zero (for instance, the second entry on the left), in which case we need to add a positive constant to satisfy the assumptions of Theorem 1. The choice of functions from Tables S1 and S2 of the OS can be guided empirically by testing procedures of the structure of space-time covariance functions as developed by Li, Genton, and Sherman (2007); extensions to the structure of cross-covariance functions are in Li, Genton, and Sherman (2008). Some examples are reported in Section 2.3.

The next result reports a modified Gneiting class, having some desirable features that will be exposed subsequently.

Theorem 2. Let $\varphi : [0, \infty) \to \mathbb{R}_+$ be completely monotone and such that $\varphi(0) = 1$. Let $\psi : [0, \infty) \to \mathbb{R}_+$ be a positive, increasing, and concave function on the positive real line. For $n \leq 3$ a positive integer, we have that

$$C(\theta, u) = \frac{\sigma^2}{\psi(|u|)^{n+2}} \varphi\{\theta\psi(|u|)\}, \quad (\theta, u) \in [0, \pi] \times \mathbb{R}, \tag{5}$$

belongs to the class $\Psi_{2n+1,T}$.

Again, the proof is deferred to Appendix A1 (OS). We call this class a *modified Gneiting class*. In the original Gneiting class, the temporal argument basically rescales the spatial one. The positive integer n in Equation (5) is required for technical reasons as detailed in Appendix A1 (OS). This class provides an easy way to generate new valid examples, again using the same ingredients and tables as illustrated for the class in Equation (4). The Gneiting class was criticized by Kent, Mohammadzadeh, and Mosammam (2011) for a possibly counterintuitive property that the authors called dimple: namely, for a given spatial lag, the temporal margins might be nonmonotonically decreasing. The authors offered conditions for the presence of a dimple under the Gneiting class. Direct inspection shows that the modified Gneiting class does not have the dimple property.

Table 1. Parametric families of members of $\Psi_{\infty,T}$ obtained through direct construction as in Theorem 3. Second column reports the analytic expression, where g is any correlation function on the real line. An additional condition is required for the fifth example, as detailed through the third column. The fourth column details the differentiability at the origin for the spatial margin. All of the members C in the second column are rescaled so that $C(0, 0) = 1$.

Family	Analytic expression	Parameters range	Differentiability at zero for $\widetilde{C}(\theta) := C(\theta, 0)$
Negative binomial	$C(\theta, u) = \left\{\frac{1-\varepsilon}{1-\varepsilon g(u)\cos\theta}\right\}^{\tau}$	$\varepsilon \in (0,1), \tau > 0$	$\widetilde{C} \in \mathcal{C}^{\infty}(\{0\})$
Multiquadric	$C(\theta, u) = \frac{(1-\varepsilon)^{2\tau}}{\{1+\varepsilon^2-2\varepsilon g(u)\cos\theta\}^{\tau}}$	$\varepsilon \in (0,1), \tau > 0$	$\widetilde{C} \in \mathcal{C}^{\infty}(\{0\})$
Sine series	$C(\theta, u) = e^{g(u)\cos\theta - 1}\{1 + g(u)\cos\theta\}/2$		$\widetilde{C} \in \mathcal{C}^{\infty}(\{0\})$
Sine power	$C(\theta, u) = 1 - 2^{-\alpha}\{1 - g(u)\cos\theta\}^{\alpha/2}$	$\alpha \in (0, 2]$	$\widetilde{C} \notin \mathcal{C}^{1}(\{0\})$ (for $\alpha \neq 2$) $\widetilde{C} \in \mathcal{C}^{\infty}(\{0\})$ (for $\alpha = 2$).
Adapted multiquadric	$C(\theta, u) = \left[\frac{\{1+g^2(u)\}(1-\varepsilon)}{1+g^2(u)-2\varepsilon g(u)\cos\theta}\right]^{\tau}$	$\varepsilon \in (0,1), \tau > 0$ $2g(\cdot)/\{1+g^2(\cdot)\}$ corr. function on $\mathbb{R}$	$\widetilde{C} \in \mathcal{C}^{\infty}(\{0\})$
Poisson	$C(\theta, u) = \exp\left[\lambda\{\cos\theta g(u) - 1\}\right]$	$\lambda > 0$	$\widetilde{C} \in \mathcal{C}^{\infty}(\{0\})$

2.2 Direct Constructions on the Sphere that are not Valid in the Euclidean Case

This section is devoted to the construction of members of the class $\Psi_{d,T}$ that cannot be adapted to Euclidean or chordal distances. The following result refers to a direct construction of members of the class $\Psi_{\infty,T}$.

Theorem 3. Let $\{g_k(\cdot)\}_{k=0}^{\infty}$ be an absolutely convergent sequence of continuous and positive definite functions on the real line, such that $g_k(0) = b_k$ for all $k = 0, 1, \ldots$, with $\{b_k\}_{k=0}^{\infty}$ being a probability mass sequence. Then,

$$C(\theta, u) = \sum_{k=0}^{\infty} g_k(u)\,(\cos\theta)^k, \qquad (\theta, u) \in [0, \pi] \times \mathbb{R}, \quad (6)$$

is a representation for members of the class $\Psi_{\infty,T}$.

The proof of a more general result is deferred to Appendix A2 (OS), where we show that Equation (6) is a special case of a construction being directly obtained by considering a Gaussian process of the type

$$Z(\mathbf{x}, t) = \sum_{k=0}^{\infty} \sum_{\nu \in \Upsilon_{k,d}} \xi_{k,\nu}(t) Y_{k,\nu,d}(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{S}^d, t \in \mathbb{R}, \quad (7)$$

where $\Upsilon_{k,d}$ is an index set specified through Appendix A2 (OS), $Y_{k,\nu,d} : \mathbb{S}^d \to \mathbb{C}$ are the normalized hyperspherical harmonics (Dai and Xu 2013), and where the set of all $\xi_{k,\nu}(t)$ forms a countable infinite sequence of Gaussian processes, with zero mean and $\mathbb{E}\xi_{k,\nu}(t)\xi_{k',\nu'}(t') = \delta_{k,k'}\delta_{\nu,\nu'}g_k(t - t')$, $t, t' \in \mathbb{R}$. Here, $\delta_{k,k'}$ denotes the usual Kronecker delta. All the technical details are deferred to Appendix A2 (OS).

Table 1 details some examples of parametric families of members of $\Psi_{\infty,T}$ obtained through direct construction. The names given to each of them (first column) are either parenthetical to the spatial margin, according to Gneiting (2013) (second and fourth entries), or to the coefficients b_k in the series expansion in Equation (6). The name "sine power" in the fourth entry might be surprising, but it comes from the fact that in this case $C(\theta, 0) = 1 - \left(\frac{\sin\theta}{2}\right)^{\alpha}$, the family introduced in Soubeyrand, Enjalbert, and Sache (2008). How to derive the examples from Table 1 is shown in Appendix A3 (OS). A relevant remark is that all the entries in Table 1 are obtained by considering $g_k(\cdot) = \varepsilon g(\cdot)^k$, for g a continuous temporal correlation function, and $\varepsilon \in (0, 1)$.

The direct construction principle seems to be a more natural solution because it directly relates to the interpretation of a process built through the series expansion as in Equation (7). At the same time, the members obtained through direct construction inherit some problems related to the Schoenberg representation of the class Ψ_d as outlined in Gneiting (2013). In particular:

- It seems that there is no way to obtain a closed form for members of Ψ_{∞}, which allows the differentiability at the origin to be indexed in a similar way as the Matérn does in Euclidean spaces. Notable attempts to solve this problem have been made in Jeong and Jun (2015) without success because closed forms are not available. Instead, the Matérn function can be composed with the chordal distance without any restriction on the parameter ν. Standard Taylor expansion shows that all of the parametric families presented in Table 1 have members whose spatial margin is infinitely differentiable at the origin, the only exception being the sine power model, for which the case $\alpha = 2$ implies that the associated function is only semipositive definite (Soubeyrand, Enjalbert, and Sache 2008).
- The adapted construction has the same problem because the Matérn covariance function can be combined with the great circle distance only for $\nu \in (0, 1/2]$. From this point of view, it might be preferable to couple the Matérn covariance with the chordal distance, which allows to recover any level of differentiability at the origin.
- The adapted construction detailed in Section 2.1 allows for rescaling the spatial component, while arguments in Theorem 8 in Gneiting (2013) imply that, in general, this cannot be done for the case presented here. This means that $C(\theta/a, u)$, for a a positive scaling parameter, is not necessarily a valid space-time covariance. Furthermore, for a fixed value of $\varepsilon \in (0, 1)$, the parameter τ in the space-time negative binomial and (adapted) multiquadric families acts as a rescaling parameter for the great circle distance.
- The direct construction in Equation (6) and examples in Table 1 allow for any type of temporal margin, provided g is a temporal correlation function. The only exception is made for the adapted multiquadric, for which the technical condition is again deferred to Appendix A3 (OS). This makes this approach more flexible than in the adapted

Gneiting class, where a strong assumption is made on the temporal margin to show that the adapted Gneiting class belongs to $\Psi_{\infty,T}$. Another limitation of the adapted and modified Gneiting classes is that they only allow for strictly positive correlations, whereas the approach based on direct construction allows for temporal correlations attaining negative values or correlations with compact support.

2.3 Some Examples and Parameterizations

This section details some examples from the adapted Gneiting class described in Section 2.1, Theorem 1, and the direct construction approach described in Section 2.2 that will be used in the analysis of the TOMS data in Section 5. The examples consider the geodesic distance on a sphere of arbitrary radius R.

Examples from the adapted Gneiting class:

- Taking the first entry in Table S1 (OS) for time and the first in Table S2 (OS) for the great circle, we obtain, after convenient parameterization (following the lines in Gneiting 2002), and after multiplication with the spatial covariance $\theta \mapsto \{1 + (\theta/c_S)\}^{-\delta}$, $c_S, \delta > 0$:

$$C(\theta, u) = \frac{\sigma^2}{\left\{1 + \left(\frac{R\theta}{c_S}\right)^{\alpha}\right\}^{\delta+\beta/2}} \times \exp\left[-\frac{\left(\frac{|u|}{c_T}\right)^{2\gamma}}{\left\{1 + \left(\frac{R\theta}{c_S}\right)^{\alpha}\right\}^{\beta\gamma}}\right], \tag{8}$$

where $c_S, c_T > 0$ are scaling parameters over space and time, respectively, the ranges of α, β, and γ are defined in the respective tables, and $\beta \in [0, 1]$ governs the separability over space and time. When $\beta = 0$, (8) reduces to the separable case. Figure S1 (OS) depicts a realization of a space-time process over the Earth ($R = 6378.88$ km), for two time instants, using a special case of the spatio-temporal covariance function in (8), under the parameterization (15) used in Section 4, and for a specified set of parameters ($c_S = 6000$ km, $c_T = 3$, and $\sigma^2 = 1$).

- Keeping the same choice for the great circle and the third choice of Table S1 (OS) for time, we obtain

$$C(\theta, u) = \frac{\sigma^2}{\left\{1 + \left(\frac{R\theta}{c_S}\right)^{\alpha}\right\}^{\delta+\beta/2}} \times \left[1 + \frac{\left(\frac{|u|}{c_T}\right)^{2\tau}}{\left\{1 + \left(\frac{R\theta}{c_S}\right)^{\alpha}\right\}^{\tau\beta}}\right]^{-\lambda}, \tag{9}$$

where the interpretation of the parameters is analogous to that in (8) and again we refer the reader to Tables S1 and S2 (OS) for the parameter space restrictions.

Examples taken from direct construction principle:

- As a first choice, from the negative binomial family (first entry in Table 1), for $\varepsilon \in (0, 1)$ and using the function $u \mapsto g(u; \alpha) := (1 + |u|^{\alpha})^{-1}$, $\alpha \in (0, 2]$, we have the following model:

$$C(\theta, u) = \sigma^2 \left[\frac{1 - \varepsilon}{1 - \varepsilon\left\{1 + \left(\frac{|u|}{c_T}\right)^{\alpha}\right\}^{-1}\cos\theta}\right]^{\tau}, \quad (\theta, u) \in [0, \pi] \times \mathbb{R}, \tag{10}$$

where σ^2 is the variance, c_T is the temporal scale, and where the positive parameter τ plays the role of spatial scaling and is analogous to the parameter c_S in the adapted Gneiting class.

- From the multiquadric (second entry in Table 1), under the same choice for the function g, we have the following model:

$$C(\theta, u) = \frac{\sigma^2(1 - \varepsilon)^{2\tau}}{\left[1 + \varepsilon^2 - 2\varepsilon\left\{1 + \left(\frac{|u|}{c_T}\right)^{\alpha}\right\}^{-1}\cos\theta\right]^{\tau}}, \quad (\theta, u) \in [0, \pi] \times \mathbb{R}, \tag{11}$$

with the same restriction on the parameters as in the previous model.

- The same choice of the function g, coupled with the sine series family (third entry in Table 1), gives the mapping

$$C(\theta, u) = \sigma^2 \upsilon \left[1 + \frac{\cos\theta}{c_S\left\{1 + \left(\frac{|u|}{c_T}\right)^{\alpha}\right\}^{-1}}\right] \times \exp\left[\frac{\cos\theta}{c_S\left\{1 + \left(\frac{|u|}{c_T}\right)^{\alpha}\right\}^{-1}}\right], \quad (\theta, u) \in [0, \pi] \times \mathbb{R}, \tag{12}$$

with c_S a positive parameter having the role of rescaling over space, and $\upsilon := 1/\{(1 + 1/c_S)\exp(1)\}$ being the normalization constant.

Figure S2 (OS) plots the covariances $C(\theta, u)$ from (12) (solid) for $\alpha = 1$, $\varepsilon = 1/2$ and from (11) (dashed) for $\alpha = \varepsilon = 1/2$, both for $\tau = 4$ and $c_T = 3$.

3. Cross-Covariance Functions of the Great Circle Distance: The Class Ψ_d^p

We now show that some parametric classes of cross-covariances proposed in the recent literature for multivariate processes can be used with the great circle distance while preserving positive definiteness over the sphere $\mathbb{S}^d$.

We start with a general result and then illustrate a very simple constructive approach that allows members of the class Ψ_{∞}^p to be obtained. Our examples put special emphasis on the case $p = 2$ for ease of illustration, but they can be extended to $p > 2$ without much mathematical effort, albeit restrictive conditions might be needed. In what follows, we write Ψ_0^p for the class of positive semidefinite matrices of real coefficients. Technical proofs are presented in Appendix B (OS).

The next result substantially rephrases a general result due to Porcu and Zastavnyi (2011) for the case of great circle distances. We omit the proof because it follows exactly the same arguments as in Porcu and Zastavnyi (2011).

Table 2. Examples of parametric members of Ψ_∞^2. Last column reports the conditions on the parameter so that the multivariate family obtained through construction in (14) belongs to Ψ_∞^2. For all cases, the cross parameters are symmetric, that is, $\lambda_{12} = \lambda_{21}$ for any parametric construction. The measures μ in the fourth column are reported from Miller and Samko (2001). In the fifth row, D_μ is the parabolic cylinder function.

Family	$\varphi_{[0,\pi]}$	Parameters	Measure $\mu(\mathrm{d}\xi)/\mathrm{d}\xi$	Conditions for Ψ_∞^2
Matérn	$(\alpha\theta)^\nu \mathcal{K}_\nu(\alpha\theta)$	$\alpha > 0, \nu \in (0, 1/2]$	$\left(\frac{\alpha^2}{4}\right)^\nu \frac{\xi^{1-\nu}}{\Gamma(\nu)} \exp(-\alpha^2/4\xi)$	see Gneiting, Kleiber, and Schlather (2010)
Generalized Cauchy	$(1 + (\theta/\beta)^a)^{-\nu}$	$\beta > 0, a \in (0, 1], \nu > 0$	$\frac{\beta^\nu}{\Gamma(\nu)} \xi^{\nu-1} \exp(-\beta\xi)$	$\beta_{12} = (\beta_{11} + \beta_{22})/2, \nu_{12} = (\nu_{11} + \nu_{22})/2$ $\rho_{12}^2 \le \frac{\beta_{11}^{\nu_{11}} \beta_{22}^{\nu_{22}}}{\beta_{12}^{2\nu_{12}}} \frac{\Gamma(\nu_{12})^2}{\Gamma(\nu_{11})\Gamma(\nu_{22})}, a_{ij} = a \in (0, 1].$
Kummer	${}_1F_1(a, c, -\theta)$	$0 < a < c$	$\frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)} \xi^{a-1}(1-t)_+^{c-a-1}$	$c_{ij} > a_{ij} > 0$ $c_{ij} = (c_{ii} + c_{jj})/2, a_{ij} = (a_{ii} + a_{jj})/2$ $\rho_{12}^2 \le \frac{\Gamma(c_{11})\Gamma(c_{22})}{\Gamma(c_{12})^2} \frac{\Gamma(a_{12})^2}{\Gamma(a_{11})\Gamma(a_{22})} \frac{\Gamma(c_{12}-a_{12})^2}{\Gamma(c_{11}-a_{11})\Gamma(c_{22}-a_{22})}$
Gauss hypergeometric	${}_1F_2(a, b, c, -\theta)$	$0 < b < c, a > 0,$	$\frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \xi^{b-1}(1-t)_+^{c-b-1}$	$c_{ij} > b_{ij} > 0, a_{ij} = a > 0$ $c_{ij} = (c_{ii} + c_{jj})/2, b_{ij} = (b_{ii} + b_{jj})/2$ $\rho_{12}^2 \le \frac{\Gamma(c_{11})\Gamma(c_{22})}{\Gamma(c_{12})^2} \frac{\Gamma(b_{12})^2}{\Gamma(b_{11})\Gamma(b_{22})} \frac{\Gamma(c_{12}-b_{12})^2}{\Gamma(c_{11}-b_{11})\Gamma(c_{22}-b_{22})}$
Parabolic cylinder	$\exp(x^2/4)D_\mu(a\theta)$	$\mu < 0, a > 0$	$\frac{a^\mu}{\Gamma(-\mu)} \exp\{-\xi^2/(2a^2)\}\xi^{-\mu-1}$	$\mu_{ij} < 0, a_{12}^2 = (a_{11}^2 + a_{22}^2)/2, \mu_{12} = (\mu_{11} + \mu_{22})/2$ $\rho_{12}^2 \le \frac{a_{11}^{\mu_{11}} a_{22}^{\mu_{22}}}{a_{12}^{2\mu_{12}}} \frac{\Gamma(-\mu_{12})^2}{\Gamma(-\mu_{11})\Gamma(-\mu_{22})}$

Theorem 4. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, $\boldsymbol{K}(\theta; \omega) : [0, \pi] \times \Omega \to \mathbb{R}^{p\times p}$, and

$$\mathbf{C}(\theta) = \int_\Omega \boldsymbol{K}(\theta; \omega)\, \mu(\mathrm{d}\omega) = \left[\int_\Omega A_{ij}(\theta; \omega)\, \mu(\mathrm{d}\omega)\right]_{i,j=1}^p, \quad \theta \in [0, \pi].$$

1. Assume that $\boldsymbol{K} = [A_{ij}(\cdot; \cdot)]_{i,j=1}^p$ satisfies the following two conditions:
   (1a) for every $i, j = 1, \ldots, p$ and $\theta \in [0, \pi]$, the function $A_{ij}(\theta; \cdot)$ belongs to $L_1(\Omega, \mathcal{F}, \mu)$;
   (1b) $\boldsymbol{K}(\cdot; \omega) \in \Psi_d^p$ for μ-almost every $\omega \in \Omega$.
   Then, $\mathbf{C} \in \Psi_d^p$.
2. Conditions (1a) and (1b) are satisfied when, for $\theta \in [0, \pi]$, $\boldsymbol{K}(\theta; \omega) = f(\theta; \omega)\boldsymbol{F}(\theta; \omega)$, where the maps $f(\theta; \omega) : [0, \pi] \times \Omega \to \mathbb{R}$ and $\boldsymbol{F}(\theta; \omega) = [F_{ij}(\theta; \omega)]_{i,j=1}^p : [0, \pi] \times \Omega \to \mathbb{R}^{p\times p}$ satisfy the following conditions:
   (a) for every $i, j = 1, \ldots, p$ and $\theta \in [0, \pi]$, the function $f(\theta; \cdot)F_{ij}(\theta; \cdot)$ belongs to $L_1(\Omega, \mathcal{F}, \mu)$;
   (b) $f(\cdot; \omega) \in \Psi_d$ for μ-almost every $\omega \in \Omega$; and
   (c) $\boldsymbol{F}(\cdot; \omega) \in \Psi_d^p$ for μ-almost every $\omega \in \Omega$.

Theorem 4 offers a construction principle that allows members of the class Ψ_∞^p to be built easily. In particular, let $\varphi(\cdot; \boldsymbol{\lambda}) : [0, \infty) \to \mathbb{R}$ be a completely monotone function with $\varphi(0; \boldsymbol{\lambda}) = 1$. Here, $\boldsymbol{\lambda} \in \mathbb{R}^m$ denotes a vector of parameters. Call $\varphi_{[0,\pi]}(\cdot; \boldsymbol{\lambda})$ the restriction of φ to the interval $[0, \pi]$. Then, the arguments in Theorem 7 in Gneiting (2013) show that $\varphi_{[0,\pi]}$ is an element of the class Ψ_∞. Now, invoking Bernstein's Theorem (Feller 1966, p. 441), we can write

$$\varphi_{[0,\pi]}(\theta; \boldsymbol{\lambda}) = \int_{[0,\infty)} \exp\{-\xi\theta\}\mu(\mathrm{d}\xi; \boldsymbol{\lambda}), \qquad \theta \in [0, \pi], \tag{13}$$

with μ being a positive and bounded measure. Observe that the integrand $\theta \mapsto \exp(-\theta\xi)$ is the restriction of a completely monotone function to the interval $[0, \pi]$ for all $\xi > 0$. Call $\boldsymbol{\mu}(\cdot) = [\mu_{ij}(\cdot)]_{i,j=1}^p$ and $\mu_{ij}(\cdot) := \mu(\cdot; \boldsymbol{\lambda}_{ij})$, for μ as defined through Equation (13). Then, we can define the mapping

$$\mathbf{C}(\theta) = \left[\sigma_i\sigma_j\rho_{ij}\mathbf{C_{ij}}(\theta)\right]_{\mathbf{i,j=1}}^{\mathbf{p}}, \qquad \theta \in [\mathbf{0}, \pi], \tag{14}$$

with ρ_{ij} a colocated correlation coefficient and $C_{ij}(\theta) = \varphi_{[0,\pi]}(\theta; \boldsymbol{\lambda}_{ij})$, $i, j = 1, \ldots, p$, and $\theta \in [0, \pi]$. Applying Theorem 4, we have that $\mathbf{C} \in \boldsymbol{\Psi}_\infty^{\mathbf{p}}$ whenever $\boldsymbol{\mu}(\xi_0)$ belongs to the class Ψ_0^p for any $\xi_0 \ge 0$.

Table 2 illustrates examples of parametric members of the class Ψ_∞^2. The conditions come from a very simple determinantal inequality of the type $\mu(\mathrm{d}\xi; \boldsymbol{\lambda}_{11})\mu(\mathrm{d}\xi; \boldsymbol{\lambda}_{22}) \ge \rho_{12}^2\mu(\mathrm{d}\xi; \boldsymbol{\lambda}_{12})^2$, $\xi \ge 0$, where for each mapping $\varphi_{[0,\pi]}$, the third column reports the associated measure in the integral representation (13) and where we set $\boldsymbol{\lambda}_{12} = \boldsymbol{\lambda}_{21}$. We omit the calculations for the examples in Table 2, since these are obtained through straightforward, albeit tedious, algebra.

Generalization to the case $p > 2$ can be achieved through diagonal dominance, a more restrictive condition than positive definiteness of $\boldsymbol{\mu}$. This often implies very restrictive conditions on the colocated correlation coefficients, as can be shown for the case of the Matérn and generalized Cauchy families, see Appendix B (OS). Observe that the multivariate Matérn family inherits the restriction on ν for the Matérn function to be completely monotone. In this respect, it might be preferable to use the chordal rather than great circle distance for the multivariate Matérn.

4. Assessing the Discrepancies Between Great Circle, Chordal, and Euclidean Map Projected Distances: A Simulation Study

We start with some details on the computation of great circle (GC), chordal (CH), and Euclidean based on map projection (MP) distances. For two location sites in longitude and latitude (expressed in decimal degrees), $P_i = (\mathrm{lon}_i, \mathrm{lat}_i)$ and $P_j = (\mathrm{lon}_j, \mathrm{lat}_j)$, and the radius R of the Earth (6378.88 km), the great

circle distance is given by $R\theta$, where

$$\theta := \theta(\mathbf{x}_i, \mathbf{x}_j) = \arccos\left(\langle \mathbf{x}_i, \mathbf{x}_j \rangle\right) = \arccos\{\sin a_i \sin a_j + \cos a_i \cos a_j \cos(b_i - b_j)\}.$$

Here $a_i = (\text{lat}_i)\pi/180$, $a_j = (\text{lat}_j)\pi/180$, $b_i = (\text{lon}_i)\pi/180$, $b_j = (\text{lon}_j)\pi/180$, $\mathbf{x}_i = (a_i, b_i)^{\mathrm{T}}$, and $\mathbf{x}_j = (a_j, b_j)^{\mathrm{T}}$. This is the natural distance to be used with global data. An approximation of such a metric is the chordal distance:

$$d_{\mathrm{CH}}(\mathbf{x}_i, \mathbf{x}_j) = R\{(\cos a_i \cos b_i - \cos a_j \cos b_j)^2 + (\cos a_i \sin b_i - \cos a_j \sin b_j)^2 + (\sin a_i - \sin a_j)^2\}^{1/2}.$$

The chordal distance obviously underestimates the GC distance, and the approximation error increases with the size of the considered portion of the Earth. Nevertheless, it has been used to analyze global data since it allows consideration of those spatial and spatio-temporal covariance models being valid on $\mathbb{R}^3$ or $\mathbb{R}^3 \times \mathbb{R}$ when the spatial argument is composed with the Euclidean distance; see, for instance, Jun and Stein (2007) or Sang, Jun, and Huang (2011).

As outlined by Banerjee (2005), when global data are available, a popular approach among GIS users is the use of the Euclidean distance obtained from a map projection, a systematic representation of all or part of the surface of the Earth on a plane. This typically comprises lines delineating meridians (longitudes) and parallels (latitudes) as required by some definitions of the projection (Banerjee 2005). Actually, there is no best projection because none of the available ones is free of distortion. Map projection is based on considering a sphere with coordinates $P_j = (\text{lon}_j, \text{lat}_j)$ and then constructing an appropriate rectangular or polar coordinate system $P(P_j) = (x_j, y_j)$ so that $x_j = f(\text{lon}_j, \text{lat}_j)$, $y_j = g(\text{lon}_j, \text{lat}_j)$, where f and g are appropriate functions to be determined based upon the properties we want our map to possess. As pointed out by Banerjee (2005), choosing the best projection can be a complicated process and largely depends on the size and location of the area under study. As an example, in our simulation study, we shall consider the so-called sinusoidal projection. This is obtained by specifying $f(\text{lon}_j, \text{lat}_j) = R\,\text{lon}_j \cos(\text{lat}_j)$ and $g(\text{lon}_j, \text{lat}_j) = R\,\text{lat}_j$.

4.1 Estimation of Space-Time Covariances on the Sphere

Assessing the discrepancies between different distance-based models in terms of estimation is a difficult task. On the one hand, models based on different metrics cannot be compared in terms of likelihood because, for two given points on the sphere, they induce different distances. On the other hand, the use of classical indicators such as the ones defined in subsequent Equations (16) and (17) imply to work with models that can be coupled with all the distances discussed in this article. Thus, we cannot make comparisons between models based on direct construction and, for instance, models that can be adapted, for example, to the chordal distance.

To illustrate how the choice of the distance affects the model estimation, we consider an increasing region (in lon/lat format) of the type $[-179.4/\kappa, 179.4/\kappa] \times [-89.5/\kappa, 89.5/\kappa]$ with $\kappa = 3, 2, 1$ (see Figure 1).

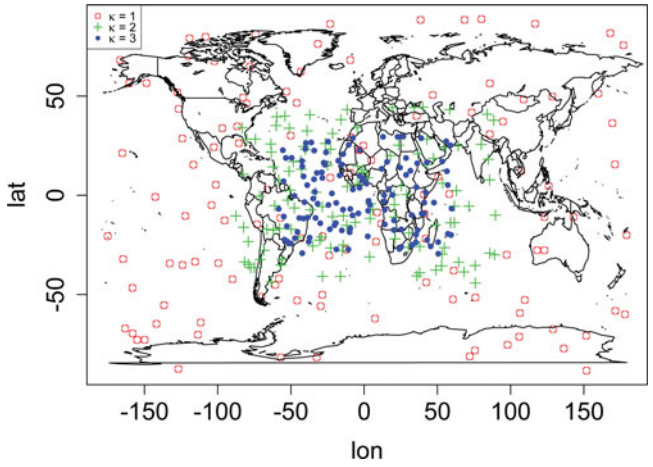

Figure 1. Portions of the Earth's surface and locations of sites considered in the simulation study, for $\kappa = 1, 2, 3$, and the portion of the Earth being (in lon/lat coordinates) equal to $[-179.4/\kappa, 179.4/\kappa] \times [-89.5/\kappa, 89.5/\kappa]$.

For each region, we simulate with Cholesky decomposition and using GC distances, 1000 realizations of a zero mean spatio-temporal Gaussian process on 120 uniformly distributed location sites, and five temporal instants, $t = 1, 2, \ldots, 5$, such that for each simulation, we have a sample of 600 observations. In particular, we use the covariance model:

$$C(\theta, u; \boldsymbol{\lambda}) = \frac{\sigma^2}{\left\{1 + \left(\frac{R\theta}{c_S}\right)\right\}} \exp\left[-\frac{|u|}{c_T\left\{1 + \left(\frac{R\theta}{c_S}\right)\right\}^{1/4}}\right], \tag{15}$$

The parameter vector $\boldsymbol{\lambda} \in \mathbb{R}^3_+$ includes c_S and c_T (respectively, spatial and temporal scale parameters) and σ^2 (the variance). This model is a special case of the model in Equation (5) under the choices $\gamma = 1/2$, $\delta = 3/4$, $\beta = 1/2$, and $\alpha = 1$. For reasons that become apparent from the exposition of our results, we use the abuse of notation $\boldsymbol{\lambda} = (c_S, c_T, \sigma^2)^{\mathrm{T}} =: ({}_1\lambda, {}_2\lambda, {}_3\lambda)^{\mathrm{T}}$.

For each region, we consider two scenarios with increasing spatial dependence ($\kappa = 1, 2, 3$):

Scenario (I) $c_T = 2$, $\sigma^2 = 1$, and $c_S = 600/\kappa$;
Scenario (II) $c_T = 2$, $\sigma^2 = 1$, and $c_S = 1, 200/\kappa$,

such that, in total, we have six cases. Note that the spatial dependence is proportional to the size of the observed region.

We estimate $\boldsymbol{\lambda}$ using maximum likelihood (ML) estimation under the model in Equation (15), using either the GC, CH, or MP distances. Thus, we use the notation $\widehat{\boldsymbol{\lambda}}^{(k)}_{\mathcal{X}}$, with $\mathcal{X}$ being either the GC or CH or MP distance, and where $k = 1, \ldots, 1000$ is the progressive number of simulations performed under a given scenario.

Figure 2 shows the boxplots of the GC versus CH and MP distances for the three regions considered in our scenarios. Apparently, CH distances are a poor approximation of the state of nature, and the discrepancy between CH and GC increases, as expected, when the size of the region increases. Note that MP distance distortion is apparent mainly in the larger region ($\kappa = 1$). Thus, we can expect a sensible difference in terms of the estimation of the spatial scale parameter.

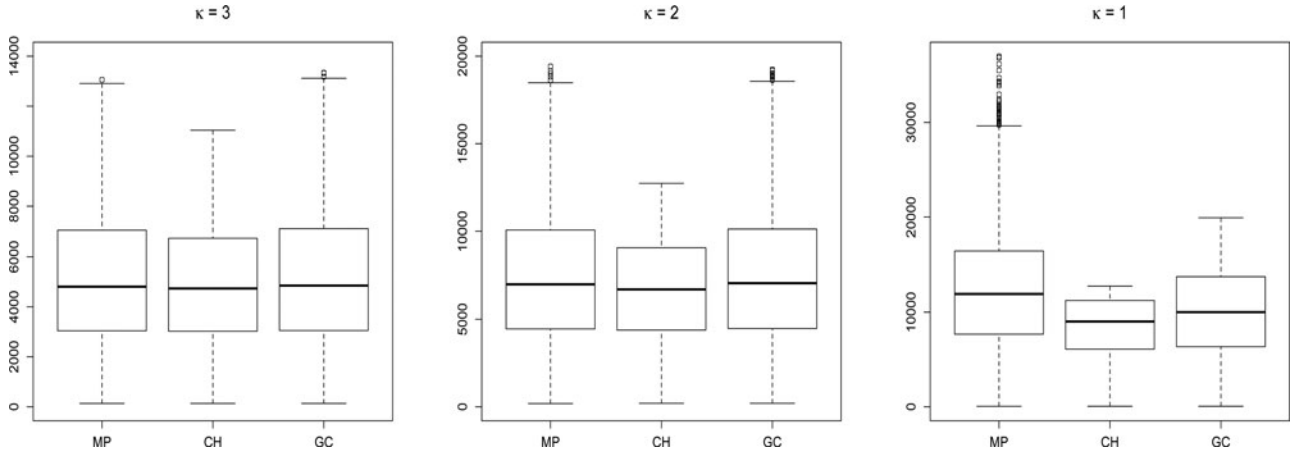

Figure 2. Boxplots of MP, CH, GC distances for the three regions considered in the simulations.

Table 3. $M(\cdot)$ as defined in (16) and $A(\cdot)$ as defined in (17) for the parameters c_S, c_T, σ^2 for the region with $\kappa = 3, 2, 1$ (3 is the smallest, 2 the medium, 1 the largest) when $c_S = 600/\kappa$ (scenario (I)) and $c_S = 1200/\kappa$ (scenario (II)).

	$\kappa = 1$		$\kappa = 2$		$\kappa = 3$	
	(I)	(II)	(I)	(II)	(I)	(II)
$M^{CH}(\hat{c}_S)$	6.25	24.07	0.97	4.06	0.42	2.19
$M^{MP}(\hat{c}_S)$	93.03	201.99	9.86	17.45	3.17	5.91
$M^{CH}(\hat{c}_T)$	0.013	0.014	0.004	0.008	0.004	0.007
$M^{MP}(\hat{c}_T)$	0.058	0.080	0.015	0.022	0.007	0.012
$M^{CH}(\hat{\sigma}^2)$	0.004	0.011	0.001	0.003	0.001	0.002
$M^{MP}(\hat{\sigma}^2)$	0.026	0.045	0.005	0.008	0.002	0.004
$A^{GC}(\hat{c}_S)$	116.47	206.77	67.74	111.52	45.67	74.90
$A^{CH}(\hat{c}_S)$	117.73	211.78	68.00	112.34	45.73	74.98
$A^{MP}(\hat{c}_S)$	133.73	261.39	68.59	112.96	45.89	74.16
$A^{GC}(\hat{c}_T)$	0.212	0.211	0.212	0.212	0.212	0.211
$A^{CH}(\hat{c}_T)$	0.212	0.212	0.212	0.212	0.212	0.212
$A^{MP}(\hat{c}_T)$	0.220	0.226	0.213	0.213	0.213	0.212
$A^{GC}(\hat{\sigma}^2)$	0.088	0.101	0.084	0.094	0.083	0.093
$A^{CH}(\hat{\sigma}^2)$	0.088	0.104	0.084	0.094	0.083	0.093
$A^{MP}(\hat{\sigma}^2)$	0.089	0.105	0.084	0.095	0.083	0.093

We now measure the discrepancy between the ML estimates using either GC (the correct distance) and CH and MP distances. Given $\widehat{\boldsymbol{\lambda}}_{\mathcal{X}}^{(k)}$, we call $M(\cdot)$ the measure

$$M^{\mathcal{X}}({}_i\widehat{\lambda}) = \sqrt{\frac{\sum_{k=1}^{1000}({}_i\widehat{\lambda}_{\mathrm{GC}}^{(k)} - {}_i\widehat{\lambda}_{\mathcal{X}}^{(k)})^2}{1000}}, \qquad i = 1, 2, 3, \quad \mathcal{X} = \mathrm{CH}, \mathrm{MP}. \tag{16}$$

We also define another measure $A(\cdot)$ by

$$A^{\mathcal{X}}({}_i\widehat{\lambda}) = \sqrt{\frac{\sum_{k=1}^{1000}({}_i\widehat{\lambda}_{\mathcal{X}}^{(k)} - {}_i\lambda)^2}{1000}}, \qquad i = 1, 2, 3, \quad \mathcal{X} = \mathrm{GC}, \mathrm{CH}, \mathrm{MP}, \tag{17}$$

where ${}_i\lambda$ denotes the nominal value chosen under one of the proposed scenarios.

The first measure compares the estimates obtained with the true distance (GC) with the estimates obtained using CH and MP distances while the second measure describes the simulated root mean square error obtained using GC, CH, and MP distances. In Table 3, we present the index M in (16) for the parameters c_S, c_T, and σ^2, as well as the index A in (17) for all possible combinations of scenarios and regions. Apparently, the choice of the type of distance does not affect the variance and temporal scale parameter estimation. For the index M, the discrepancy between ML estimation with GC distance and CH or MP distances increases, as expected, as the size of the region of observation increases. Moreover, this discrepancy increases when the strength of the spatial correlation increases. This fact is apparent when using the MP distance.

As for index A, the results in Table 3 suggest that GC distance is the best choice from a statistical efficiency viewpoint. As expected, the parameter that is most affected by the choice of the distance is the spatial scale while the variance and the temporal scale are only slightly affected. The estimates based on the CH distance perform well enough when compared to GC while the estimates based on the MP distance present more variability particularly when considering the larger region ($\kappa = 1$).

4.2 Estimation of Spatial Cross-Covariances on the Sphere

We replicate this experiment now considering a bivariate spatial setting. Specifically, we consider a bivariate exponential model defined by

$$C_{ij}(\theta; \boldsymbol{\lambda}) = \sigma_i \sigma_j \rho_{ij} \exp\left(-\frac{R\theta}{c_{ij}}\right), \quad \rho_{ii} = 1, \quad i, j = 1, 2, \tag{18}$$

obtained by setting $p = 2$ and $\nu_{ij} = 1/2$ in Equation (13) OS, and, to simplify the estimation task, we consider the constraints $c_{12} = (c_{11} + c_{22})/2$. Thus, in this case, $\boldsymbol{\lambda} = (\sigma_1, \sigma_2, \rho_{12}, c_{11}, c_{22})^{\mathrm{T}}$, and we consider 1000 simulations obtained with Cholesky decomposition over the regions defined in Figure 1 ($\kappa = 3, 2, 1$) under the following scenarios:

Scenario (III) $\rho_{12} = 0.2, \sigma_1 = 1, \sigma_2 = 1, c_{11} = 600/\kappa$, and $c_{22} = 720/\kappa$;

Scenario (IV) $\rho_{12} = 0.2$, $\sigma_1 = 1$, $\sigma_2 = 1$, $c_{11} = 1200/\kappa$, and $c_{22} = 1440/\kappa$.

The estimation is performed with maximum likelihood, and we compute the measures M and A as previously defined. The results are reported in Table 4, and the conclusions are basically the same: only the spatial scale parameters' estimation is

Table 4. $M(\cdot)$ as defined in (16) and $A(\cdot)$ as defined in (17) for the parameters a_1, a_2, σ_1, σ_2, and ρ_{12} for the region with $\kappa = 3, 2, 1$ (3 is the smallest, 2 the medium, 1 the largest) when $a_1 = 600/\kappa$, $a = 2 = 720/\kappa$ (scenario (III)) and $a_1 = 1200/\kappa$, $a_2 = 1440/\kappa$ (scenario (IV)).

	$\kappa = 1$		$\kappa = 2$		$\kappa = 3$	
	(III)	(IV)	(III)	(IV)	(III)	(IV)
$M^{CH}(\hat{c}_{11})$	70.38	135.64	30.15	54.65	12.71	31.90
$M^{MP}(\hat{c}_{11})$	204.74	221.80	52.47	63.12	20.09	36.62
$M^{CH}(\hat{c}_{22})$	71.15	189.37	29.40	73.96	9.60	38.92
$M^{MP}(\hat{c}_{22})$	204.54	270.78	46.51	82.01	21.02	41.73
$M^{CH}(\hat{\sigma}_1)$	0.010	0.022	0.004	0.012	0.003	0.010
$M^{MP}(\hat{\sigma}_1)$	0.053	0.034	0.006	0.013	0.004	0.011
$M^{CH}(\hat{\sigma}_2)$	0.010	0.030	0.004	0.017	0.003	0.013
$M^{MP}(\hat{\sigma}_2)$	0.053	0.044	0.006	0.018	0.004	0.014
$M^{CH}(\hat{\rho}_{12})$	0.004	0.004	0.004	0.003	0.003	0.003
$M^{MP}(\hat{\rho}_{12})$	0.058	0.034	0.006	0.009	0.004	0.005
$A^{GC}(\hat{c}_{11})$	176.20	252.92	127.37	149.83	88.89	102.29
$A^{CH}(\hat{c}_{11})$	178.64	262.06	126.08	148.90	88.35	102.78
$A^{MP}(\hat{c}_{11})$	201.87	293.79	129.40	153.51	88.93	102.35
$A^{GC}(\hat{c}_{21})$	184.18	280.23	126.64	161.76	87.47	114.35
$A^{CH}(\hat{c}_{22})$	183.94	283.93	129.62	161.49	87.43	112.45
$A^{MP}(\hat{c}_{22})$	213.09	307.07	128.48	161.97	87.88	113.72
$A^{GC}(\hat{\sigma}_1)$	0.067	0.076	0.066	0.072	0.066	0.071
$A^{CH}(\hat{\sigma}_1)$	0.067	0.077	0.066	0.072	0.066	0.071
$A^{MP}(\hat{\sigma}_1)$	0.070	0.082	0.066	0.072	0.066	0.071
$A^{GC}(\hat{\sigma}_2)$	0.069	0.075	0.066	0.072	0.066	0.073
$A^{CH}(\hat{\sigma}_2)$	0.069	0.075	0.066	0.072	0.066	0.074
$A^{MP}(\hat{\sigma}_2)$	0.071	0.082	0.067	0.072	0.066	0.072
$A^{GC}(\hat{\rho}_{12})$	0.089	0.089	0.089	0.089	0.089	0.089
$A^{CH}(\hat{\rho}_{12})$	0.089	0.089	0.089	0.089	0.089	0.089
$A^{MP}(\hat{\rho}_{12})$	0.091	0.096	0.089	0.089	0.088	0.089

affected by the choice of the distance. In particular, the estimation based on CH and MP distances perform badly when the size of the region of observation ($\kappa = 3, 2, 1$) and/or the strength of the dependence (moving from scenario (III) to scenario (IV)) is increased.

5. TOMS Data

This section analyzes Level-3 Total Ozone Mapping Spectrometer (TOMS) data, which include daily total column ozone levels. The data are located on a spatially regular grid (1° latitude by 1.25° longitude away from the poles) as reported by Jun and Stein (2008), and we refer to their article for a detailed description of the data. The original data range from latitude interval $[-89.5, 89.5]$ to the longitudes $[-180, 180]$. Jun and Stein (2008) considered a single spatial realization over the horizontal window $[-89.5, 89.5]$ (with all longitudes) whereas we consider the daily realizations over 2 weeks (15 observations in time) and select the horizontal window $[-35, 35]$ and all longitudes as well. In total, we have a regular grid of 20,160 points (288 longitudinal and 70 latitudinal) observed during 15 days, for a total of 302,400 observations.

We have missing data and follow the same procedure proposed by Jun and Stein (2008) to handle them. Namely, for each missing data point, we naively put the average data obtained from eight neighboring cells in space and two contiguous temporal instants, so that we use 24 observations for each local averaging. The available dataset is definitely too large for likelihood estimation. We therefore select a subset of the grid with 336 spatial points and all temporal observations, for a total of 5040 observations. We then detrend the data using spatio-temporal splines and consider the residuals as a realization from a zero mean space-time Gaussian random field. We consider three classes of covariance models:

A. Two models based on the adapted Gneiting classes as in Theorem 1, which can be coupled with any of GC, CH, or MP. Specifically, we used:
   A.1 The model in Equation (8) with $\alpha = 2, \gamma = 1/2$, and $\delta + \beta/2 = 1$;
   A.2 The model in Equation (9) with $\alpha = 2, \tau = 1/2, \lambda = 1$, and $\delta + \beta/2 = 1$;
B. Three models based on direct construction, hence valid with GC only. Specifically:
   B.1 The model in Equation (12);
   B.2 The model in Equation (11) with $\varepsilon = 0.226$;
   B.3 The model in Equation (10) with $\varepsilon = 0.961$.
   To choose ε, we first maximize the likelihood with respect to all the parameters in Equations (11) and (10), respectively. Then, we take the value of ε, which maximizes the likelihood and maximizes the likelihood again with respect to the other parameters.
C. A model based on the Gneiting class valid using CH and MP distances:
   C.1 The model in eq. (16) in Gneiting (2002) is valid if coupled with CH or MP (but not with GC), with $\nu = 1.5$, $\alpha = 1$, and $\delta + \beta/2 = 1$. This kind of model is a nonseparable covariance with spatial margin of the Matérn type, where we fix the smoothness parameter equal to 1.5.

We also tried other models, such as those coming from the modified Gneiting class, but these were considerably outperformed by the ones proposed in this section. Thus, we do not report the results concerning their constructions here.

For all of the selected models, we consider a space-time nugget effect, to account for potential microscale effects. This in turn induces discontinuities at the origin of the associated covariance functions. For all the parametric families described above, the parameters are estimated through maximum likelihood techniques.

For all of the models we need to estimate four parameters: in classes A and C, the variance (denoted σ^2), the spatial scale (c_S), the temporal scale (c_T), and the nonseparability parameter (β), whereas in class B, the variance (σ^2), the spatial scale (τ), the temporal scale (c_T), and the temporal smoothing parameter (α).

Table 5 summarizes our findings: for obvious reasons, we only compare the performance (based on likelihood) in terms of estimation for models based on the same distance. Such a comparison is mistakenly reported, for instance, in Jeong and Jun (2015). In light of this, apparently the models based on direct construction outperform the others coupled with GC distance (especially model B.3). Among the models based on CH distance, clearly model C.1 outperforms the others. As a simple diagnostic, models B.3 and C.1 are compared in Figure S4 (OS) in terms of their fit to the marginal spatial and temporal empirical semivariograms. It seems that model B.3 offers a better fit to the temporal empirical variogram, whereas the behaviors of the spatial margins are very similar.

Table 5. Parameter estimates for the TOMS data analyzed in Section 5 for covariance models A (with great circle (GC), chordal (CH), and map projection (MP) distances) and for covariance models B (with great circle (GC) only) and for covariance models C (with great circle (CH) only).

Distance	GC	CH	MP	GC	CH	MP	GC	CH	MP
Model		A.1			A.2			C.1	
c_S	742.7	743.9.3	734.2	681.3	672.2	733.3	–	450.8	417.5
c_T	2.54	2.54	2.17	1.76	1.67	1.71	–	1.56	1.29
β	1	1	0.95	1	1	1	–	0.95	0.89
σ^2	102.6	102.1	102.7	106.1	103.0	111.4	–	97.9	98.5
Nugget	9.81	9.82	6.35	6.30	5.60	4.97	–	12.26	5.72
Likelihood	− 17233.8	− 17234.0	− 17296.5	− 17257.1	− 17258.5	− 17317.3	–	− 17156.3	− 17223.9
Model		B.1			B.2			B.3	
τ	0.01	–	–	144.09	–	–	4.04	–	–
c_T	51.05	–	–	50.58	–	–	36.85	–	–
α	1.49	–	–	1.49	–	–	1.61	–	–
σ^2	85.8	–	–	86.0	–	–	89.9	–	–
Nugget	18.60	–	–	18.54	–	–	16.84	–	–
Likelihood	− 17168.3	–	–	− 17167.9	–	–	− 17162.8	–	–

We also assess the prediction performance of the covariance models proposed in our analysis, through predictive scores (see Table S4 in OS). This allow us to (a) compare the predictive skills of each estimated covariance model and (b) compare, when allowable, for a fixed estimated covariance model, the impact on prediction when using different distances. For the first goal, it can be appreciated that the best models from a prediction point of view are the models C.1 and the model B.3, followed by the models B.1, B.2, and the models of the class A. In particular the model C.1 based on the chordal distance slightly outperforms the model B.3 based on direct construction. For the second goal, the comparison is possible only for the covariance models in the class A, and we see that, as expected from the simulation study, there is not much difference in terms of prediction when comparing the predictive scores using GC or CH distances while the use of the MP distances affects considerably the quality of the prediction.

Summing up, a first conclusion is that all our models favor a nonseparable structure over the sphere cross time. This is directly highlighted by the value of the beta parameter in the adapted Gneiting class, while models based on direct construction do not admit separability as special case. Second, the models based on direct construction outperform the models belonging to the adapted Gneiting class in terms of prediction. Nevertheless, the Gneiting model based on chordal distance, with a Matérn spatial margin, slightly outperforms the other models. This might be explained by the fact that models based on direct construction do not allow the smoothness of the underlying Gaussian field to be parameterized.

The analysis and the simulation study have been carried out using an upcoming version of the R package CompRandFld (Padoan and Bevilacqua 2015), avalaible at http://cran.r-project.org/.

6. Discussion

In this article, we have proposed stationary covariance functions for processes evolving temporally over the sphere, as well as cross-covariance functions for multivariate random fields defined over the sphere. We have illustrated several methods of construction and provided closed-form expressions for both cases. A simulation study assessed the discrepancy between the great circle, chordal, and Euclidean from map projection distances both in terms of estimation and prediction in a space-time and a bivariate spatial settings, where the space is in this case the Earth. The results highlight that when global data are available, the choice of the chordal or the Euclidean distance based on map projection can seriously affect the quality of the estimation particularly when data are available for a large portion of the Earth's surface. In this sense, the classes of parametric covariance models that are valid on the sphere presented in this article are useful when working with space-time or multivariate data. The predictions are less affected by the choice of distance, but the GC distance is nevertheless the most realistic from a physics point of view. In particular, prediction results based on GC and CH distances are very similar while the use of the MP distance can affect the quality of the prediction. We have revisited the analysis of TOMS data and investigated differences in terms of estimation and prediction between great circle, chordal, and Euclidean from map projection distance-based approaches.

The extension of our work to nonstationary models can be tackled by using spatially adaptive parameters in the completely monotone function in Equation (4). One possibility is to follow the lines of the proof proposed by Porcu, Mateu, and Christakos (2010). An intriguing challenge is to extend the approach by Jun and Stein (2008) to generate nonstationary models on the sphere. Another possible generalization is to make the anisotropic covariance models of Hitczenko and Stein (2012) on the sphere functions of the great circle distance. Some challenges regarding the class $\Psi^p_{d,T}$ remain open. One of them is to find the analog of Schoenberg's representation in terms of Gegenbauer polynomials for the elements of such classes.

Supplementary Materials

The supplementary materials contain: Tables of completely monotone and Bernstein functions; a realization of a space-time Gaussian process with covariance from the adapted Gneiting class; plots of space-time covariances from the direct construction; the results of the simulation study and the TOMS data in terms of prediction; and all proofs of the results of the paper.

Acknowledgment

The authors are grateful to Christian Berg and to Morten Nielsen for checking the proofs of the results, and for very interesting discussions.

Funding

Research work of Emilio Porcu was partially supported by grant FONDECYT 1130647 from the Chilean government. Research work of Moreno Bevilacqua was partially supported by grant FONDECYT 11121408 from the Chilean government. The work of Marc G. Genton was supported by King Abdullah University of Science and Technology (KAUST).

References

Apanasovich, T. V., and Genton, M. G. (2010), "Cross-Covariance Functions for Multivariate Random Fields Based on Latent Dimensions," *Biometrika*, 97, 15–30. [889]

Apanasovich, T. V., Genton, M. G., and Sun, Y. (2012), "A Valid Matérn Class of Cross-Covariance Functions for Multivariate Random Fields With Any Number of Components," *Journal of the American Statistical Association*, 107, 180–193. [889]

Banerjee, S. (2005), "On Geodetic Distance Computations in Spatial Modeling," *Biometrics*, 61, 617–625. [888,894]

Dai, F., and Xu, Y. (2013), *Approximation Theory and Harmonic Analysis on Spheres and Balls (Springer Monographs in Mathematics)*, New York: Springer. [891]

Feller, W. (1966), *An Introduction to Probability Theory and Its Applications* (Vol. II), New York: Wiley. [893]

Furrer, R., and Genton, M. G. (2011), "Aggregation-Cokriging for Highly-Multivariate Spatial Data," *Biometrika*, 98, 615–631. [889]

Genton, M. G., and Kleiber, W. (2015), "Cross-Covariance Functions for Multivariate Geostatistics" (with discussion), *Statistical Science*, 30, 147–163. [889]

Gneiting, T. (2002), "Nonseparable, Stationary Covariance Functions for Space-Time Data," *Journal of the American Statistical Association*, 97, 590–600. [889,890]

——— (2013), "Strictly and Non-Strictly Positive Definite Functions on Spheres," *Bernoulli*, 19, 1327–1349. [888,889,890,891,893]

Gneiting, T., Genton, M. G., and Guttorp, P. (2007), "Geostatistical Space-Time Models, Stationarity, Separability and Full Symmetry," in *Statistics of Spatio-Temporal Systems (Monographs in Statistics and Applied Probability)*, eds. B. Finkenstaedt, L. Held, and V. Isham, Boca Raton, FL: Chapman & Hall/CRC Press, pp. 151–175. [890]

Gneiting, T., Kleiber, W., and Schlather, M. (2010), "Matérn Cross-Covariance Functions for Multivariate Random Fields," *Journal of the American Statistical Association*, 105, 1167–1177. [889,893]

Hitczenko, M., and Stein, M. L. (2012), "Some Theory for Anisotropic Processes on the Sphere," *Statistical Methodology*, 9, 211–227. [897]

Jeong, J., and Jun, M. (2015), "A Class of Matérn-like Covariance Functions for Smooth Processes on a Sphere," *Spatial Statistics*, 11, 1–18. [896]

Jun, M. (2011), "Non-Stationary Cross-Covariance Models for Multivariate Processes on a Globe," *Scandinavian Journal of Statistics*, 38, 726–747. [889]

Jun, M., and Stein, M. L. (2007), "An Approach to Producing Space-Time Covariance Functions on Spheres," *Technometrics*, 49, 468–479. [888,894]

——— (2008), "Nonstationary Covariance Models for Global Data," *Annals of Applied Statistics*, 2, 1271–1289. [888,889,890,896,897]

Kent, J. T., Mohammadzadeh, M., and Mosammam, A. M. (2011), "The Dimple in Gneiting's Spatial-Temporal Covariance Model," *Biometrika*, 98, 489–494. [889,890]

Kleiber, W., and Genton, M. G. (2013), "Spatially Varying Cross-Correlation Coefficients in the Presence of Nugget Effects," *Biometrika*, 100, 213–220. [889]

Li, B., Genton, M. G., and Sherman, M. (2007), "A Nonparametric Assessment of Properties of Space-Time Covariance Functions," *Journal of the American Statistical Association*, 102, 736–744. [890]

——— (2008), "Testing the Covariance Structure of Multivariate Random Fields," *Biometrika*, 95, 813–829. [890]

Miller, K. S., and Samko, S. G. (2001), "Completely Monotonic Functions," *Integral Transforms and Special Functions*, 12, 389–402. [890]

North, G. R., Wang, J., and Genton, M. G. (2011), "Correlation Models for Temperature Fields," *Journal of Climate*, 24, 5850–5862. [888]

Padoan, S., and Bevilacqua, M. (2015), "Analysis of Random Fields Using CompRandFld," *Journal of Statistical Software*, 63, 1–27. [897]

Porcu, E., Mateu, J., and Christakos, G. (2010), "Quasi-Arithmetic Means of Covariance Functions with Potential Applications to Space-Time Data," *Journal of Multivariate Analysis*, 100, 1830–1844. [897]

Porcu, E., and Schilling, R. (2011), "From Schoenberg to Pick-Nevanlinna: Toward a Complete Picture of the Variogram Class," *Bernoulli*, 17, 441–455. [890]

Porcu, E., and Zastavnyi, V. (2011), "Characterization Theorems for Some Classes of Covariance Functions Associated to Vector Valued Random Fields," *Journal of Multivariate Analysis*, 102, 1293–1301. [892]

Sang, H., Jun, M., and Huang, J. Z. (2011), "Covariance Approximation for Large Multivariate Spatial Data Sets with an Application to Multiple Climate Model Errors," *Annals of Applied Statistics*, 5, 2519–2548. [894]

Schoenberg, I. J. (1938), "Metric Spaces and Completely Monotone Functions," *Annals of Mathematics*, 39, 811–841. [888]

Soubeyrand, S., Enjalbert, J., and Sache, I. (2008), "Accounting for Roughness of Circular Processes: Using Gaussian Random Processes to Model the Anisotropic Spread of Airborne Plant Disease," *Theoretical Population Biology*, 73, 92–103. [891]

Stein, M. L. (1999), *Statistical Interpolation of Spatial Data: Some Theory for Kriging*, New York: Springer. [889]

Yadrenko, M. I. (1983), *Spectral Theory of Random Fields (Translation Series in Mathematics and Engineering)*, New York: Optimization Software. [889,890]

Zastavnyi, V. P., and Porcu, E. (2011), "Characterization Theorems for the Gneiting Class of Space-Time Covariances," *Bernoulli*, 17, 456–465. [890]

Zhang, H. (2007), "Maximum-Likelihood Estimation for Multivariate Spatial Linear Coregionalization Models," *Environmetrics*, 18, 125–139. [889]