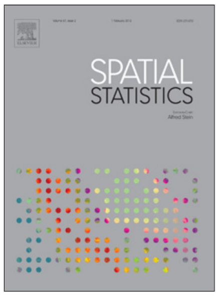
Journal Pre-proof



Families of covariance functions for bivariate random fields on spheres

Moreno Bevilacqua, Peter Diggle, Emilio Porcu

PII: S2211-6753(20)30042-7
DOI: https://doi.org/10.1016/j.spasta.2020.100448
Reference: SPASTA 100448

To appear in: *Spatial Statistics*

Received date : 24 January 2019
Revised date : 15 April 2020
Accepted date : 15 April 2020

Please cite this article as: M. Bevilacqua, P. Diggle and E. Porcu, Families of covariance functions for bivariate random fields on spheres. *Spatial Statistics* (2020), doi: https://doi.org/10.1016/j.spasta.2020.100448.

Families of Covariance Functions for Bivariate Random Fields on Spheres

Moreno Bevilacqua,[1] Peter Diggle,[2] and Emilio Porcu [3]

Abstract

This paper proposes a new class of covariance functions for bivariate random fields on spheres, having the same properties as the bivariate Matérn model proposed in Euclidean spaces. The new class depends on the geodesic distance on a sphere; it allows for indexing differentiability (in the mean square sense) and fractal dimensions of the components of any bivariate Gaussian random field having such covariance structure. We find parameter conditions ensuring positive definiteness. We discuss other possible models and illustrate our findings through a simulation study, where we explore the performance of maximum likelihood estimation method for the parameters of the new covariance function. A data illustration then follows, through a bivariate data set of temperatures and precipitations, observed over a large portion of the Earth, provided by the National Oceanic and Atmospheric Administration Earth System Research Laboratory.

Keywords, GREAT-CIRCLE DISTANCE; CROSS CORRELATION; $\mathcal{F}$ CLASS; MATÉRN CLASS; MEAN SQUARE DIFFERENTIABILITY

[1]Department of Statistics, University of Valparaiso, Valparaiso, Chile.
Millennium Nucleus Center for the Discovery of Structures in Complex Data , Chile
[2]CHICAS, Lancaster Medical School, Lancaster University,
Lancaster, LA1 4YB, United Kingdom.
[3]School of Computer Science and Statistics, Trinity College at Dublin
& School Mathematics, Statistics, and Physics, Newcastle University
& Millennium Nucleus Center for the Discovery of Structures in Complex Data , Chile
E-mail: georgepolya01@gmail.com, porcue@tcd.ie.

1 Introduction

1.1 Context

The paper deals with modeling, inference and prediction for bivariate Gaussian random fields defined on the unit sphere $\mathbb{S}^2 = \{\boldsymbol{x} \in \mathbb{R}^3, \ \|\boldsymbol{x}\| = 1\}$. The reason for the interest in this geometry is given by the increasing availability of multivariate data collected over the whole planet, or a big portion of it. For instance, monitoring several georeferenced variables is a common practice in a wide range of disciplines such as climatology and oceanography (Reinsel et al., 1981; Di Lorenzo et al., 2014; Nychka et al., 2015; Combes et al., 2017; Edwards et al., 2019).

Our approach considers the observations as the partial realization of a bivariate Gaussian random field, denoted as $\boldsymbol{Z} = \{\boldsymbol{Z}(\boldsymbol{x}) = (Z_1(\boldsymbol{x}), Z_2(\boldsymbol{x}))^\top : \boldsymbol{x} \in \mathbb{S}^2\}$, where $\top$ is the transpose operator. Gaussianity assumption plays a central role in many scientific fields such as atmospheric, environmental and geological sciences, and provides a building block for non-Gaussian random fields (Alegría et al., 2017). The components Z_i, $i = 1, 2$, for the vector $\boldsymbol{Z}$ are called scalar random fields.
Let $\theta(\cdot, \cdot) : \mathbb{S}^2 \times \mathbb{S}^2 \to [0, \pi]$ be the geodesic distance, defined as

$$\theta(\boldsymbol{x}, \boldsymbol{y}) = \arccos(\boldsymbol{x}^\top \boldsymbol{y}), \qquad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{S}^2.$$

We focus on the analysis of geodesically isotropic (Porcu et al., 2016; Alegría et al., 2019) covariance functions $\boldsymbol{C} : [0, \pi] \to \mathbb{R}^{2\times 2}$, being matrix valued mappings, whose elements are defined as $\mathrm{C}_{ij}(\theta(\boldsymbol{x}, \boldsymbol{y})) = \mathrm{cov}\{Z_i(\boldsymbol{x}), Z_j(\boldsymbol{y})\}$, $i, j = 1, 2$. For the reminder of the paper, we always assume pointwise continuity for the elements C_{ij} of the matrix-valued mapping $\boldsymbol{C}$. Also, we use θ instead of $\theta(\boldsymbol{x}, \boldsymbol{y})$ for simplicity.

The mapping $\boldsymbol{C}(\theta)$ must be positive definite, which means that

$$\sum_{\ell=1}^{n} \sum_{r=1}^{n} \boldsymbol{a}_\ell^\top \boldsymbol{C}(\theta(\boldsymbol{x}_\ell, \boldsymbol{x}_r)) \boldsymbol{a}_r \geq 0, \tag{1}$$

for all positive integer n, $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} \subset \mathbb{S}^2$ and $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\} \subset \mathbb{R}^2$. An alternative modeling strategy on the function $\mathbf{C}$ might be based on the use of the chordal distance, being an approximation of the geodesic distance θ. The constructive criticism expressed in Banerjee (2005) , Gneiting (2013) and Porcu et al. (2016) on the use of such a metric motivates us to build models that depend exclusively on geodesic distance. Alternatively, one might consider models based on Euclidean distances coupled with some map projection, but such models have been shown by Porcu et al. (2018), through simulation, to be outperformed by models based on geodesic distance in terms of maximum likelihood (ML) estimation of the scale parameter.

1.2 Literature Review

Multivariate covariance functions in Euclidean spaces have become ubiquitous and we refer the reader to Genton and Kleiber (2015) for a detailed account. Yet, the literature on multivariate covariance models on spheres has been sparse, with the exceptions of Porcu et al. (2016), Alegría and Porcu (2017) and Alegría et al. (2019).

Some construction principles might be adapted from Euclidean spaces. Alegría et al. (2019) give an account on how to adapt methods on Euclidean spaces to the sphere. For instance, the linear model of coregionalization (Wackernagel, 2003) is based on representing any component of the bivariate field $\boldsymbol{Z}$, as a linear combination of latent, uncorrelated fields. The constructive criticisms in Gneiting et al. (2010) and Daley et al. (2015) motivate us not to propose this model on the sphere. For instance, the smoothness of any component of the multivariate field amounts to that of the roughest underlying univariate process. Moreover, the number of parameters can quickly become massive as the number of components increases. Instead, one might resort to the scale mixture techniques proposed in Porcu and Zastavnyi (2011) to create bivariate covariance functions. Some examples will be provided subsequently. Latent dimension approaches (Porcu et al., 2006; Apanasovich and Genton, 2010; Porcu and Zastavnyi, 2011) might also be easily adapted to the sphere. Finally, Alegría et al. (2019) call the following construction principle *multivariate parametric adaptation*: let p be a positive integer. Let $\{C(\cdot;\boldsymbol{\lambda}) : [0,\pi] \to \mathbb{R},\ \boldsymbol{\lambda} \in \mathbb{R}^p\}$ be a parametric family of geodesically isotropic univariate correlation functions ($C(0;\boldsymbol{\lambda}) = 1$) indexed by a parameter vector $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_p)^\top$. Call

$\boldsymbol{\lambda}_{ij} = (\lambda_{ij,1}, \ldots, \lambda_{ij,p})^\top$, $i,j = 1,2$ four parameter vectors in $\mathbb{R}^p$. Then, define $\boldsymbol{C} : [0,\pi] \to \mathbb{R}^{2\times 2}$ through

$$\boldsymbol{C}(\theta) = [C_{ij}(\theta)]_{i,j=1}^{2}, \qquad \theta \in [0,\pi], \qquad i,j = 1,2,$$

with elements C_{ij} defined as

$$C_{ij}(\theta) = \sigma_{ii}\sigma_{jj}\rho_{ij} C(\theta; \boldsymbol{\lambda}_{ij}), \qquad \theta \in [0,\pi], \qquad i,j = 1,2, \tag{2}$$

where σ_{ii}^2 is that variance of the ith component of the bivariate random field, where ρ_{12} is the colocated correlation coefficient. Thus, we can write ρ_{ij} in the equation above because $\rho_{ii} = 1$ by construction. In Euclidean spaces this strategy has been adopted by Gneiting et al. (2010), Apanasovich et al. (2012) and by Daley et al. (2015).

Let $\boldsymbol{\lambda} = (\alpha, \nu)^\top$, with α and ν being strictly positive. Then, Gneiting et al. (2010) have coupled the construction (2) with the Matérn family, so that

$$C_{ij}(\theta) = \sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{M}_{\alpha_{ij},\nu_{ij}}(\theta) = \sigma_{ii}\sigma_{jj}\rho_{ij}\frac{2^{1-\nu_{ij}}}{\Gamma(\nu_{ij})}\left(\frac{\theta}{\alpha_{ij}}\right)^{\nu_{ij}} \mathcal{K}_{\nu_{ij}}\left(\frac{\theta}{\alpha_{ij}}\right), \qquad \theta \in [0,\pi], \tag{3}$$

where α_{ij} are scaling parameters and where the parameters ν_{ij} index differentiability at the origin, and consequently the differentiability, in the mean square sense, of the associated Gaussian random field. Here, $\mathcal{K}_\nu$ is the McDonald function (Abramowitz and Stegun, 1970). Unfortunately, Theorem 7 in Gneiting (2013) implies that, in the univariate case (a scalar valued random field) the function in Equation (3) is positive definite only under a very severe restriction on the smoothing parameter ν ($\nu \in (0, 1/2]$), implying that the Matérn class on the sphere is valid only for the case of very rough processes. This results in a considerable drawback when modeling spatial data on the sphere. Apparently, any bivariate structure of the type (2) coupled with a Matérn choice would inherit such a restriction.

An alternative model can be obtained by coupling the construction (2) with the Generalized Wendland class, $\mathcal{W}_{\alpha,\nu,\mu}$ (see Bevilacqua et al., 2019, with the references therein, for a recent account) defined, for $\nu = 0$, as

$$\mathcal{W}_{\alpha,0,\mu}(\theta) = \left(1 - \frac{\theta}{\alpha}\right)_+^{\mu}, \qquad \theta \in [0,\pi], \tag{4}$$

and, for $\nu > 0$, as

$$\mathcal{W}_{\alpha,\nu,\mu}(\theta) = \frac{1}{B(2\nu, \mu+1)} \int_{\theta/\alpha}^{1} u \left(u^2 - \left(\frac{\theta}{\alpha}\right)^2\right)^{\nu-1} \mathcal{W}_{\alpha,0,\mu}(u)\,\mathrm{d}u, \qquad \theta < \alpha, \qquad \theta \in [0,\pi] \tag{5}$$

where $0 < \alpha < \pi$ implies that the covariance is compactly supported and $(x)_+$ denotes the positive part of the real number x. For both cases, μ is strictly positive. The parameter $\nu \geq 0$ has the same role as the smoothness parameter in the Matérn class and the compact support can lead to considerable computational benefits when handling the associated sparse covariance matrix (Furrer et al., 2006). Positive definiteness of the Generalized Wendland class on $\mathbb{S}^d$ is an open problem. For the special cases $\nu = k$, a nonnegative integer, positive definiteness is guaranteed when $\mu > k+2$ on a three dimensional sphere (and, consequently, on $\mathbb{S}^2$; see Gneiting, 2013). In this case $\mathcal{W}_{\alpha,k,\mu}$ can be written as:

$$\mathcal{W}_{\alpha,k,\mu}(\theta) = \left(1 - \frac{\theta}{\alpha}\right)_+^{k+\mu} \mathcal{P}_k\left(\frac{\theta}{\alpha}\right), \qquad \theta \in [0,\pi], \qquad k = 0,1,2,\ldots \tag{6}$$

where $\mathcal{P}_k$ is polynomial of degree at most k. The Wendland class has been used as a radial function in d-dimensional Euclidean spaces $\mathbb{R}^d$: Wendland (1995) showed that the polynomial degree k is minimal for given space dimension d and smoothness $2k$. Daley et al. (2015) have coupled the construction (2) with the family (6), that is $C_{ij}(\theta) = \sigma_{ii}\sigma_{jj}\rho_{ij}C(\theta; \boldsymbol{\lambda}_{ij}) = \sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{W}_{\alpha_{ij},\nu_{ij},\mu_{ij}}(\theta)$, $\theta \in [0,\pi]$.

1.3 Motivation and Our Contribution

The severe restriction of the smoothing parameter ν in the Matérn class results in a very limited appeal for a bivariate construction (2) based on Matérn functions. For univariate random fields, this has already been put as an open problem by Porcu et al. (2018), and then solved by Alegría et al. (2018), who proposed the so called $\mathcal{F} = \mathcal{F}_{\tau,\alpha,\nu}$ family of functions, defined through the identity

$$\mathcal{F}_{\tau,\alpha,\nu}(\theta) = \frac{B(\alpha, \nu+\tau)}{B(\alpha,\nu)} {}_2F_1(\tau, \alpha, \alpha+\nu+\tau; \cos\theta), \qquad \theta \in [0,\pi], \tag{7}$$

where τ, α and ν are strictly positive parameters and where ${}_2F_1$ is the Gauss Hypergeometric function

$${}_2F_1(a,b,c;x) = \sum_{k=0}^{\infty} \frac{(a)_k (b)_k}{(c)_k} \frac{x^k}{k!}, \quad |x| < 1,$$

with $(\cdot)_k$ being the Pochhammer symbol. Finally, $B(\cdot,\cdot)$ is the Beta function (Abramowitz and Stegun, 1970). Alegría et al. (2018) show that (7) is k times differentiable at the origin if and only if $\lfloor \nu/2 \rfloor > k$, for any τ, α, and where $\lfloor x \rfloor$ denotes the largest integer being smaller than $x \in \mathbb{R}$. Further, they show that the $\mathcal{F}$ family allows to index fractal dimension through the parameter ν. Also, a wealth of special cases that can be written in closed form is available. Thus, the $\mathcal{F}$ class shares all the properties of the Matérn class on planar surfaces and becomes a reference for spatial analysis of isotropic Gaussian fields on spheres. Further, Alegría et al. (2018) show how to extend this isotropic construction to an axially symmetric covariance, a natural property for the analysis of climate data, where nonstationarities over latitude are typically encountered.

This paper proposes bivariate covariance models based on the multivariate parametric adaptation (2) coupled with the choice $C_{ij}(\cdot) = \sigma_{ii}\sigma_{jj}\rho_{ij}C(\cdot;\boldsymbol{\lambda}_{ij}) = \sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{F}_{\tau_{ij},\alpha_{ij},\nu_{ij}}$, the $\mathcal{F}$ family in Equation (7), for $\boldsymbol{\lambda} = (\tau, \alpha, \nu)^\top$. Practical parameterizations will be discussed as well. A simulation study in Section 3 explores the performance of the ML estimation method when estimating the involved parameters. Further, Section 4 illustrates a bivariate data set of temperature and precipitation, observed on a large portion of the Earth, using the bivariate $\mathcal{F}$ model. The paper concludes with discussion.

2 Results

We start with the main result of the paper, which shows the positive definiteness of the bivariate construction proposed subsequently. Some notation is needed. For any positive integer p and vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^p$, we write $\boldsymbol{a} = \boldsymbol{b}$ to denote that $a_k = b_k$, $k = 1, \ldots, p$.

Theorem 2.1 *Let $\boldsymbol{\lambda}_{ij} = (\alpha_{ij}, \nu_{ij}, \tau_{ij})^\top$, $i, j = 1, 2$ be vectors of strictly positive parameters. Consider the model in Equation (2), with*

$$C_{ij}(\theta) = \sigma_{ii}\sigma_{jj}\rho_{ij}C(\cdot;\boldsymbol{\lambda}_{ij}) = \sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{F}_{\tau_{ij},\alpha_{ij},\nu_{ij}}(\theta), \qquad \theta \in [0, \pi], \tag{8}$$

with $\boldsymbol{\lambda}_{12} = \boldsymbol{\lambda}_{21}$*. Let* $\tau_{12} \le \min\{\tau_{11}, \tau_{22}\}$*. If, either,*

(A) *Scale mixture based conditions:*

1. $\alpha_{12} = \frac{1}{2}(\alpha_{11} + \alpha_{22})$;
2. $\nu_{11} + \nu_{22} \le 2\nu_{12}$;
3. $|\rho_{12}| \le \frac{B(\alpha_{12},\nu_{12})}{(B(\alpha_{11},\nu_{11})B(\alpha_{22},\nu_{22}))^{1/2}}$,

or

(B) *Spectral conditions:*

1. $\alpha_{12} \le \min\{\alpha_{11}, \alpha_{22}\}$;
2. $\nu_{12} \ge \max\{\nu_{11} + (\alpha_{11} - \alpha_{12}) + (\tau_{11} - \tau_{12}), \nu_{22} + (\alpha_{22} - \alpha_{12}) + (\tau_{22} - \tau_{12})\}$;
3. $|\rho_{12}| \le \frac{B(\alpha_{12},\nu_{12})}{(B(\alpha_{11},\nu_{11})B(\alpha_{22},\nu_{22}))^{1/2}} \frac{B(\alpha_{11},\nu_{11}+\tau_{11})^{1/2} B(\alpha_{22},\nu_{22}+\tau_{22})^{1/2}}{B(\alpha_{12},\nu_{12}+\tau_{12})}$.

Then, the mapping $\boldsymbol{C} : [0, \pi] \to \mathbb{R}^{2\times 2}$ *with elements* C_{ij} *as in (8), is a geodesically isotropic matrix–valued covariance mapping associated with a Gaussian random field* $\boldsymbol{Z}$ *defined over* $\mathbb{S}^2$*.*

2.1 Other Models

In the following we detail on how to construct, within the framework of multivariate parametric adaptation, alternative models to the bivariate $\mathcal{F}$ family. Lemma A.2 proves that, for $\widetilde{\boldsymbol{\lambda}}_{ij} = (\delta_{ij}, \tau_{ij})^\top$ and for the model in Equation (2), with

$$C_{ij}(\theta) = C(\cdot; \widetilde{\boldsymbol{\lambda}}_{ij}) = \sigma_{ii}\sigma_{jj}\widetilde{\rho}_{ij}\mathcal{N}_{\delta_{ij},\tau_{ij}}(\theta), \qquad \theta \in [0, \pi],$$

with $\widetilde{\boldsymbol{\lambda}}_{ij} = \widetilde{\boldsymbol{\lambda}}_{ji}$ and equality intended as pointwise, and with $\mathcal{N}_{\delta,\tau}$ defined as

$$\mathcal{N}_{\delta,\tau}(\theta) = \left(\frac{1-\delta}{1-\delta\cos\theta}\right)^{\tau}, \qquad \theta \in [0, \pi], \tag{9}$$

positive definiteness is attained provided

1. $\tau_{12} \leq \min(\tau_{11}, \tau_{22})$ and $\delta_{12} \leq \min(\delta_{11}, \delta_{22})$;

2. $|\widetilde{\rho}_{12}| \leq \frac{(1-\delta_{11})^{\tau_{11}/2}(1-\delta_{22})^{\tau_{22}/2}}{(1-\delta_{12})^{\tau_{12}}}$.

Other models can be adapted from Euclidean spaces by mimicking the scale mixture in Lemma A.3:

$$C_{ij}(\theta) = \sigma_{ii}\sigma_{jj}\rho_{ij}\int_0^\infty C(\theta;\xi)G_{ij}(\xi)\mathrm{d}\xi, \qquad \in [0,\pi], \tag{10}$$

where $C(\cdot;\xi)$ is positive definite on $\mathbb{S}^2$ for all ξ, and where the matrix $\boldsymbol{G}(\xi)$ having elements $g_{ij}(\xi)$ is positive definite for any fixed ξ. Some technical conditions might be needed for (10) to be well defined. To make some examples,

(I) A Bivariate Wendland structure. Let $\mathcal{W}_{\alpha,k,\mu}$ as defined in Equation (6). Then,

$$\boldsymbol{C}(\theta) = \left[\sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{W}_{\alpha_{ij},k,\mu_{ij}}(\theta),\right]_{i,j=1}^2, \qquad \theta \in [0,\pi],$$

is obtained by considering the parametric restrictions in Theorem 1 of Daley et al. (2015). Notice that when $k=0$, then $\mathcal{P}_0$ in Equation (6) is identically equal to 1. For $k=1$, $\mathcal{P}_1(\theta) = (1+(\mu+1)\theta/\alpha)$.

(II) A Bivariate Exponential Model. Using (I) with the model $\mathcal{W}_{\alpha,0,\mu}$ we have that the model

$$\boldsymbol{C}(\theta) = \left[\sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{W}_{n/\alpha_{ij},0,n}(\theta),\right]_{i,j=1}^2, \qquad \theta \in [0,\pi],$$

for n a positive integer, is a valid model under the relevant conditions as in Theorem 1 of Daley et al. (2015). Hence, the model

$$\boldsymbol{C}(\theta) = \left[\sigma_{ii}\sigma_{jj}\rho_{ij}\lim_{n\to\infty}\mathcal{W}_{n/\alpha_{ij},0,n}(\theta),\right]_{i,j=1}^2 = \left[\sigma_{ii}\sigma_{jj}\rho_{ij}\exp(-\alpha_{ij}\theta),\right]_{i,j=1}^2, \qquad \theta \in [0,\pi],$$

is a valid model under the constraints on the parameters α_{ij} coming from Theorem 1 of Daley et al. (2015).

Porcu et al. (2016) have proposed other bivariate models based on scale mixtures. None of the examples proposed there (see their table 2) allow to index differentiability at the origin and/or fractal dimension. Further, some parametric forms are valid only under some severe restriction on the parameters.

Other models might be obtained on the basis of spectral conditions. For instance, let us consider the sine power model (Soubeyrand et al., 2008), defined by

$$\mathcal{S}_\alpha(\theta) = 1 - \left(\sin\frac{\theta}{2}\right)^\alpha, \qquad \theta \in [0, \pi],$$

where $\alpha \in (0, 2]$ (for $\alpha = 2$ the model is semi-positive definite only). Such a model is non differentiable at the origin when $\alpha \in (0, 2)$ and infinitely differentiable when $\alpha = 2$. Using the arguments in Appendix A.2 in Soubeyrand et al. (2008) in concert with the proof of Assertion **(B)** of Theorem 2.1, it is easy to show that the bivariate model with elements

$$C_{ij}(\theta) = \sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{S}_{\alpha_{ij}}(\theta), \qquad \theta \in [0, \pi],$$

is positive definite if $\alpha_{12} \geq \max\{\alpha_{11}, \alpha_{22}\}$ and $\rho_{12}^2 \leq \alpha_{11}\alpha_{22}/\alpha_{12}^2$.

2.2 Practical Choices and Parameterizations

Alegría et al. (2018) have shown some identifiability issue with the $\mathcal{F}$ class. Such a problem can be circumvented through a different parameterization. We follow Alegría et al. (2018) and consider the parameterization

$$\boldsymbol{C}(\theta) = [\sigma_i\sigma_j\rho_{ij}\mathcal{F}_{1/\alpha_{ij},1/\alpha_{ij}+0.5,\nu_{ij}}(\theta)]_{i,j=1}^2, \tag{11}$$

which allows us to identify the parameters $\alpha_{ij} > 0$ as correlation ranges. The conditions for positive definiteness in Theorem 2.1 are then changed to, respectively:

(A) Scale mixture conditions:

1. $\alpha_{12} = 2\frac{\alpha_{11}\alpha_{22}}{\alpha_{11}+\alpha_{22}}$;

2. $\nu_{11} + \nu_{22} \leq 2\nu_{12}$;
3. $|\rho_{12}| \leq \frac{B(1/\alpha_{12}+1/2,\nu_{12})}{\left(B(1/\alpha_{11}+1/2,\nu_{11})B(1/\alpha_{22}+1/2,\nu_{22})\right)^{1/2}}$.

(B) Spectral Conditions:

1. $\alpha_{12} \geq \max(\alpha_{11}, \alpha_{22})$;
2. $\nu_{12} \geq \max\{2(1/\alpha_{11} - 1/\alpha_{12}) + \nu_{11}, 2(1/\alpha_{22} - 1/\alpha_{12}) + \nu_{22}\}$;
3. $|\rho_{12}| \leq \frac{B(\frac{1}{\alpha_{12}}+\frac{1}{2},\nu_{12})}{B(\frac{1}{\alpha_{12}}+\frac{1}{2},\frac{1}{\alpha_{12}}+\nu_{12})} \times \left(\frac{B(\frac{1}{\alpha_{11}}+\frac{1}{2},\frac{1}{\alpha_{11}}+\nu_{11})B(\frac{1}{\alpha_{22}}+\frac{1}{2},\frac{1}{\alpha_{22}}+\nu_{22})}{B(\frac{1}{\alpha_{11}}+\frac{1}{2},\nu_{11})B(\frac{1}{\alpha_{22}}+\frac{1}{2},\nu_{22})}\right)^{\frac{1}{2}}$.

Note that when $\alpha_{11} = \alpha_{22} = \alpha_{12}$ and $\nu_{11} = \nu_{22} = \nu_{12}$ then in both cases the condition on the colocated parameter is $|\rho_{12}| \leq 1$, as expected.

3 Simulation Study

The main goal of this section is to analyze the performance of the ML method for the bivariate $\mathcal{F}$ covariance model estimation. Let $\boldsymbol{x}_i$, $i = 1, \ldots, N$ be distinct points of $\mathbb{S}^2$, set $\boldsymbol{Z}_{k;N} = (Z_k(\boldsymbol{x}_1), \ldots Z_k(\boldsymbol{x}_N))^\top$, $k = 1, 2$ and let $\boldsymbol{Z}_N = (\boldsymbol{Z}_{1;N}^\top, \boldsymbol{Z}_{2;N}^\top)^\top$ be a partial realization from a zero mean bivariate Gaussian random field with bivariate $\mathcal{F}$ covariance model, under the parameterization given in Section 2.2.

The log-likelihood, up to an additive constant, can be written as

$$l_N(\boldsymbol{\phi}) = -\frac{1}{2}\log|\boldsymbol{C}_N(\boldsymbol{\phi})| - \frac{1}{2}\boldsymbol{Z}_N^\top[\boldsymbol{C}_N(\boldsymbol{\phi})]^{-1}\boldsymbol{Z}_N, \tag{12}$$

where $\boldsymbol{C}_N(\boldsymbol{\phi}) := [\boldsymbol{C}_{ij}]_{i,j=1}^2$ with $\boldsymbol{C}_{ij} = \left[\sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{F}_{1/\alpha_{ij},1/\alpha_{ij}+0.5,\nu_{ij}}(\theta(\boldsymbol{x}_l, \boldsymbol{x}_m))\right]_{l,m=1}^N$ and $\boldsymbol{\phi} = (\sigma_1^2, \sigma_2^2, \rho_{12}, \alpha, \nu)^\top$ or $\boldsymbol{\phi} = (\sigma_1^2, \sigma_2^2, \rho_{12}, \alpha_{11}, \alpha_{22}, \alpha_{12}, \nu_{11}, \nu_{22}, \nu_{12})^\top$ depending on whether a separable (*i.e.*, $\nu = \nu_{11} = \nu_{22} = \nu_{12}$ and $\alpha = \alpha_{11} = \alpha_{22} = \alpha_{12}$) or a nonseparable bivariate $\mathcal{F}$ covariance model is considered.

Scenario **I** considers $N = 200$ points being uniformly distributed on the unit sphere, and sets $\sigma_1^2 = \sigma_2^2 = 1$, $\alpha = 0.3$, $\rho_{12} = 0.1, 0.4, 0.7$ and $\nu = 0.5, 2.5$. This setting corresponds to a continuous, non-differentiable ($\nu = 0.5$) and once differentiable ($\nu = 2.5$) bivariate random field respectively, with increasing positive correlation between the components of the bivariate Gaussian field. We simulate, with Cholesky decomposition, 500 realizations and perform ML estimation by maximizing the function (12) with respect to $\boldsymbol{\phi}$. Table 1 depicts bias and mean square error for each parameter, and Figure 1 reports the centered boxplots of the ML estimates when $\rho_{12} = 0.4$.

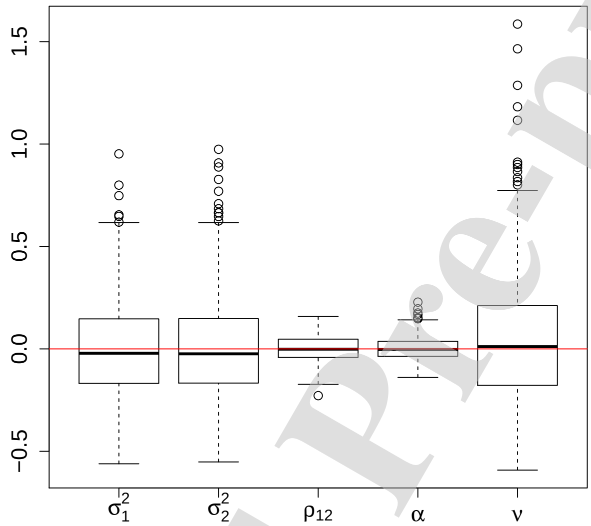


Figure 1: Centered boxplots of ML estimates, under Scenario **I**, when $\sigma_1^2 = 1$, $\sigma_2^2 = 1$, $\rho_{12} = 0.4$, $\alpha = 0.3$, and $\nu = 2.5$ (from left to right).

Increasing the smoothness parameter leads to a larger variability for the variances parameters and a smaller variability for the scale parameters. Moreover, the variability of the colocated correlation parameter is not affected by the values of the smoothness parameters. This is consistent with the results in Bevilacqua et al. (2015) which show that the asymptotic variance of the ML estimator of ρ_{12}, under increasing domain asymptotics, does not depend on spatial distance irrespectively of the type of bivariate model considered. Finally, increasing ρ_{12} does not affect the variability of both variance and scale parameters.

Scenario **II** considers the same number and location of points on the sphere for Scenario **I**. We consider a nonseparable $\mathcal{F}$ covariance model by fixing $\sigma_1^2 = \sigma_2^2 = 1$, $\alpha_{11} = 0.3$, $\alpha_{22} = 0.28$, $\alpha_{12} = 0.3$, $\nu_1 = 0.5$, $\nu_2 = 2.5$, $\nu_{12} = 3.1$ and $\rho_{12} = 0.2$. Marginal and cross-covariances are shown in Figure 3 (a). Note that, under this setting, parameters match the spectral conditions **(B)** given in Section 2.2.

According to the mixing conditions **(A)** in Section 2.2, we consider another parameter setting by fixing $\sigma_1^2 = \sigma_2^2 = 1$, $\alpha_{11} = 0.25$, $\alpha_{22} = 0.28$, $\alpha_{12} = 0.264$, $\nu_1 = 0.5$, $\nu_2 = 2.5$, $\nu_{12} = 1.5$ and $\rho_{12} = 0.5$.

For both cases, we obtain a bivariate random field with the first and second component being non differentiable, and once mean square differentiable, respectively. This can be clearly appreciated from Figure 3 (b) and (c) where a partial realization defined on a portion of the planet Earth of a bivariate Gaussian random field, under the first setting of Scenario **II**, is depicted.

For both settings of Scenario **II**, we simulate with Cholesky decomposition 500 realizations, and we perform ML estimation maximizing the function (12) with respect to $\boldsymbol{\phi}$, while keeping fixed the cross scale and cross smoothness parameters α_{12} and ν_{12}. Table 2 depicts bias and mean square error for each parameter under both settings and Figure 2 reports the centered boxplots of the ML estimates under the second setting. The estimates are unbiased, and first setting shows larger variability than the second setting, owing to a different magnitude of the spatial scale parameters.

The simulations and ML estimates of the bivariate $\mathcal{F}$ model have been performed using an upcoming version of the R package `GeoModels` (Bevilacqua and Morales-Oñate, 2018).

4 Data Illustration

In this section, we apply the proposed bivariate $\mathcal{F}$ model to a data set of temperature and precipitation. The original dataset contains measurements of monthly means (from January 1948) of surface air temperatures and precipitable water content from the National Oceanic and Atmospheric Adminis-

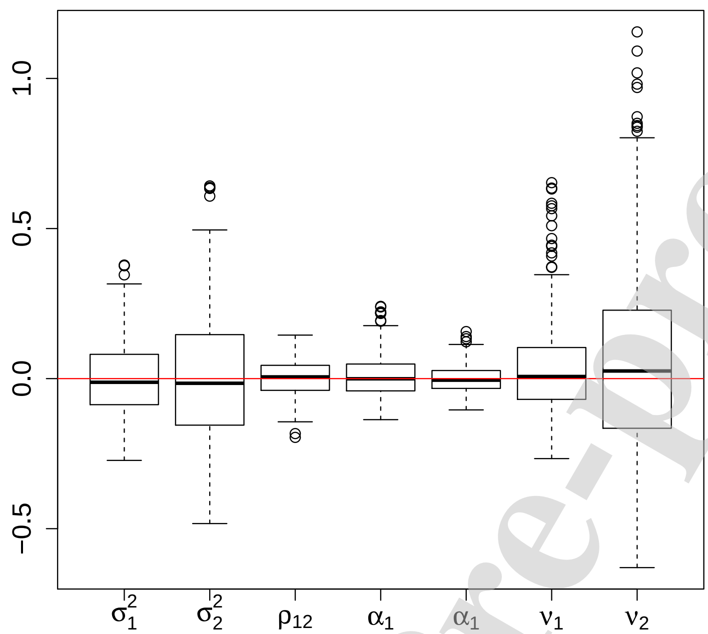


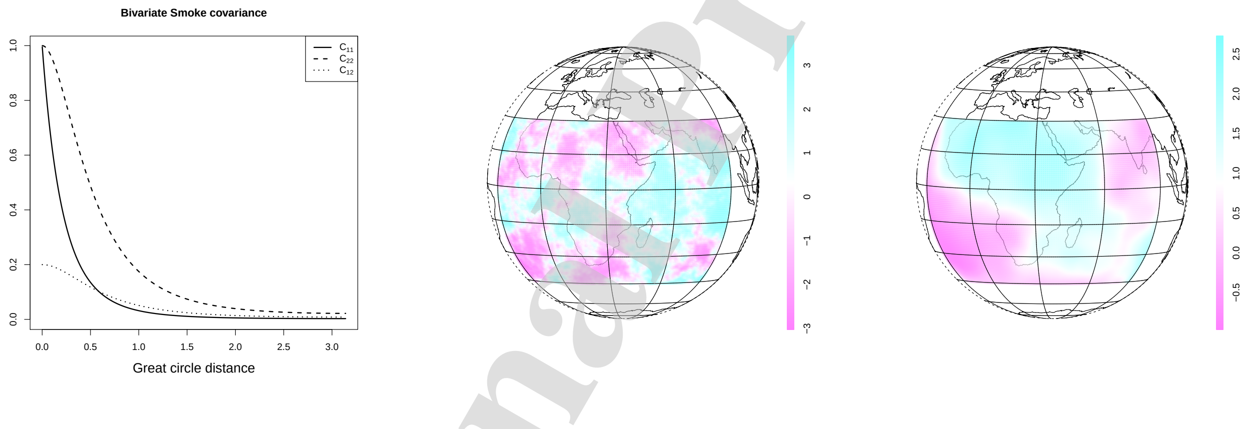
Figure 2: Centered boxplots of ML estimates, under Scenario **II**, when $\sigma_1^2 = 1$, $\sigma_2^2 = 1$, $\rho_{12} = 0.5$, $\alpha_{11} = 0.25$, $\alpha_{22} = 0.28$, $\nu_{11} = 0.5$ and $\nu_{22} = 2.5$ (from left to right).

	$\rho_{12} = 0.1$				$\rho_{12} = 0.4$				$\rho_{12} = 0.7$			
	$\nu = 0.5$		$\nu = 2.5$		$\nu = 0.5$		$\nu = 2.5$		$\nu = 0.5$		$\nu = 2.5$	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
$\hat{\sigma}_1^2$	−0.00537	0.01741	0.00021	0.05721	−0.00535	0.01740	0.00026	0.05725	−0.00533	0.01740	0.00002	0.05721
$\hat{\sigma}_2^2$	−0.00068	0.01765	0.00552	0.06145	−0.00015	0.01834	0.00678	0.06274	−0.00100	0.01893	0.00615	0.06283
$\hat{\rho}_{12}$	0.00107	0.00543	0.00144	0.00537	−0.00051	0.00396	−0.00014	0.00391	−0.00125	0.00147	−0.00103	0.00145
$\hat{\alpha}$	0.01129	0.00724	0.00254	0.00318	0.01128	0.00723	0.00254	0.00318	0.01128	0.00723	0.00253	0.00318
$\hat{\nu}$	0.01174	0.01082	0.04282	0.10681	0.01180	0.01082	0.04283	0.10687	0.01185	0.01087	0.04290	0.10688

Table 1: Bias and MSE when estimating with ML a separable bivariate $\mathcal{F}$ model (Scenario **I**). True parameters values are $\sigma_1^2 = \sigma_1^2 = 1$, $\alpha = 0.3$, $\nu = 0.5, 2.5$, $\rho_{12} = 0.1, 0.4, 0.7$.

	1		2	
	Bias	MSE	Bias	MSE
$\hat{\sigma}_1^2$	0.00426	0.02300	0.00042	0.01457
$\hat{\sigma}_2^2$	0.02408	0.08106	−0.00105	0.04101
$\hat{\rho}_{12}$	−0.01427	0.01556	0.00293	0.00352
$\hat{\alpha}_1$	0.02205	0.01488	0.00763	0.00481
$\hat{\alpha}_2$	0.00785	0.00467	−0.00044	0.00202
$\hat{\nu}_1$	0.03044	0.03097	0.02763	0.02191
$\hat{\nu}_2$	0.04697	0.17188	0.05296	0.09732

Table 2: Bias and MSE when estimating with ML a nonseparable bivariate $\mathcal{F}$ model under the first and second setting of Scenario **II**. True parameters of first setting are $\sigma_1^2 = \sigma_2^2 = 1$, $\alpha_{11} = 0.3$, $\alpha_{22} = 0.28$, $\alpha_{12} = 0.3$, $\nu_1 = 0.5$, $\nu_2 = 2.5$, $\nu_{12} = 3.1$ and $\rho_{12} = 0.2$. True parameters of second setting are $\sigma_1^2 = \sigma_2^2 = 1$, $\alpha_{11} = 0.25$, $\alpha_{22} = 0.28$, $\alpha_{12} = 0.264$, $\nu_1 = 0.5$, $\nu_2 = 2.5$, $\nu_{12} = 1.5$ and $\rho_{12} = 0.5$.



(a) (b) (c)

Figure 3: From left to right: the bivariate $\mathcal{F}$ model with $\sigma_1^2 = \sigma_2^2 = 1$, $\rho_{12} = 0.2$, $\alpha_{11} = 0.3$, $\alpha_{22} = 0.28$, $\alpha_{12} = 0.3$ and $\nu_1 = 0.5$, $\nu_2 = 2.5$ $\nu_{12} = 3.1$. A realization of a bivariate Gaussian random field on the planet Earth with the bivariate covariance models depicted in (a).

tration (NOAA) Earth System Research Laboratory (data are downloaded from *www.esrl.noaa.gov*). Temperatures are measured in Kelvin degrees and precipitable water content is in kg/m^2. Both variables are observed with a spatial resolution of 2.5×2.5 degrees of longitude and latitude. To make ML estimation feasible, we select a subset of the data, and in particular we focus on December 2006 and in the region with longitudes between 50 and 150 degrees of latitudes between -50 and 0 degrees. The resulting dataset consists of 506 observations (253 for each variable).

Following Li and Zhang (2011), we first detrend the data using splines to remove the cyclic pattern of both variables along the longitude and latitude directions, and then regard the residuals as a realization from a zero mean bivariate Gaussian random field. In Figure 4, we show the boxplots of both the original data set and the associated residuals for the temperature and precipitation data, in terms of latitudes. It becomes apparent how the detrending technique alleviates considerably the effect of latitude on data. Moreover, Figure 4 depicts the normal quantile-quantile plot of the residuals for the temperature and precipitation data. The assumption of Gaussianity in both cases seems quite reasonable.

We consider three bivariate separable models from the $\mathcal{F}$, Wendland, and Matérn classes. Specifically, we consider two separable models ($\mathcal{F}$ and Wendland) using the great circle distance, θ, and a separable Matérn model using the chordal distance defined as $d_{\rm CH} = 2\sin(\theta/2)$, that is (note: $\mathcal{B}$ stands for bivariate and S stands for separable):

$$\mathcal{B}\mathcal{W}^S(\theta) = [\sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{W}_{\alpha,6,\mu}(\theta)]_{i,j=1}^2,$$

$$\mathcal{B}\mathcal{F}^S(\theta) = [\sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{F}_{1/\alpha,1/\alpha+0.5,\nu}(\theta)]_{i,j=1}^2,$$

with $\theta \in [0,\pi]$, and

$$\mathcal{B}\mathcal{M}^S(d_{\rm CH}) = [\sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{M}_{\alpha,\nu}(d_{\rm CH})]_{i,j=1}^2, \qquad d_{\rm CH} \in [0,2].$$

The choice of the chordal distance in a bivariate Matérn model allows to have no restrictions on the smoothing parameter ν (see Porcu et al., 2018, for a recent account).

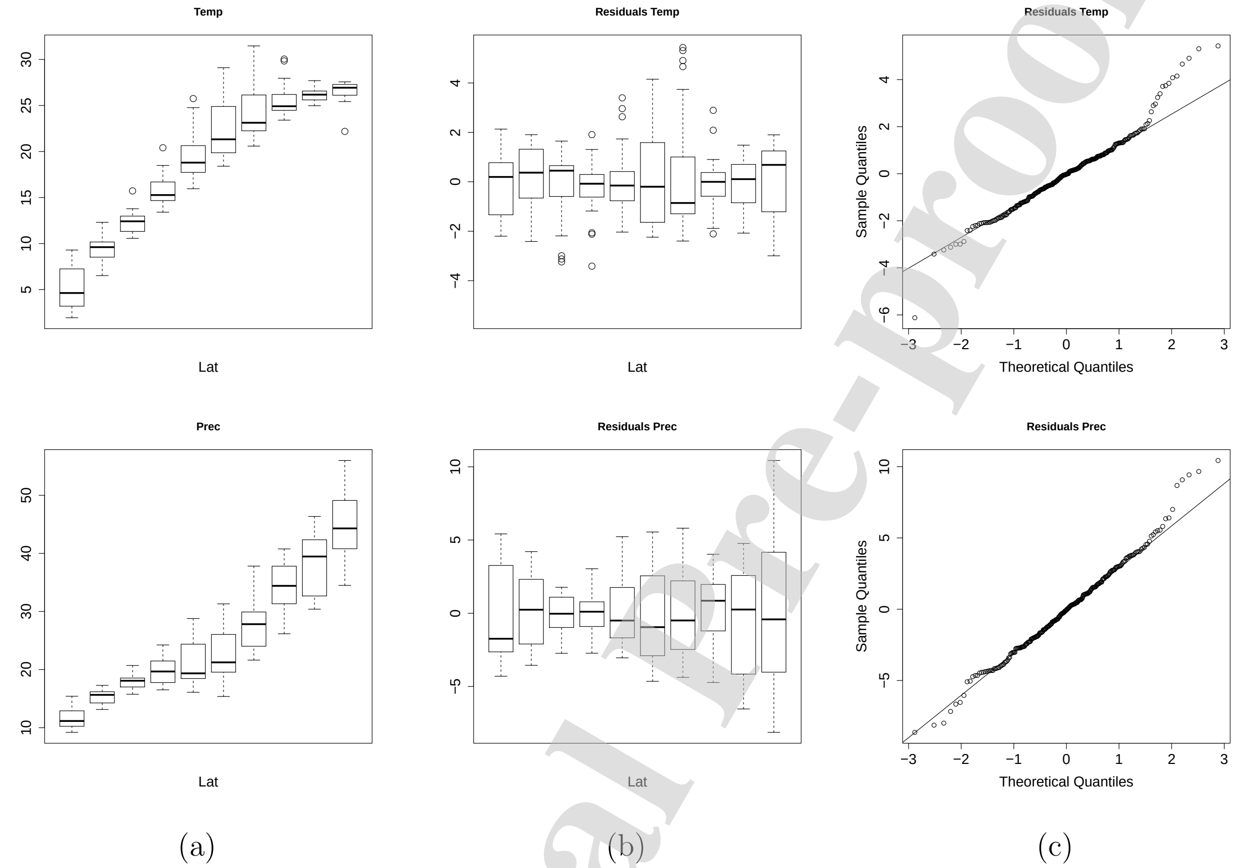


Figure 4: Top part (from left to right): boxplots by latitudes for original Temperature data, boxplots by latitudes for the associated residuals and Gaussian quantile-quantile plot for the residuals. Bottom part (from left to right): the same graphical representations as in the top part for the Precipitation variable.

We also consider the three models in their nonseparable versions (NS stands for nonseparable here):

$$\mathcal{BW}^{NS}(\theta) = [\sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{W}_{\alpha_{ij},6,\mu_{ij}}(\theta)]_{i,j=1}^{2},$$

$$\mathcal{BF}^{NS}(\theta) = [\sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{F}_{1/\alpha_{ij},1/\alpha_{ij}+0.5,\nu_{ij}}(\theta)]_{i,j=1}^{2},$$

with $\theta \in [0, \pi]$, and

$$\mathcal{BM}^{NS}(d_{CH}) = [\sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{M}_{\alpha_{ij},\nu_{ij}}(d_{CH})]_{i,j=1}^{2}, \qquad d_{CH} \in [0, 2].$$

Table 3 reports the comparison between the six considered bivariate models fits and the associated ML estimates. For both separable and nonseparable cases, the $\mathcal{F}$ model achieves the largest value of the likelihood and, as expected, the nonseparable models clearly outperform the associated separable version. Although the bivariate $\mathcal{F}$ and the bivariate Matérn are not comparable in terms of likelihood (the models are based on different metrics), we can appreciate that the parameters of the $\mathcal{F}$ and the Matérn are rather comparable in terms of scale, smoothness and colocated correlation parameter. Note that $\hat{\nu}_{12} < \hat{\nu}_{ii}$, $i = 1, 2$ and this condition does not match with the spectral or the mixing conditions given in Section 2.2. Nevertheless, the nonseparable $\mathcal{F}$ model is still valid under the setting of the obtained ML estimates. This fact is not surprising, since the conditions given in Section 2.2 are sufficient only.

We further evaluate the predictive performances of the different Gaussian bivariate random fields using RMSE and MAE indicators (see Zhang and Wang, 2010, and the references therein). Specifically, we use the following re-sampling approach: we randomly choose 200 location sites and we use the estimates in order to compute, for each bivariate covariance model, RMSE and MAE values at the remaining 53 spatial locations using cokriging predictor for each variable. Specifically, we condider

$$\text{RMSE}_X = \sqrt{\frac{1}{53}\sum_{k=1}^{53}(Z_X(\boldsymbol{s}_k) - \widehat{Z}_X(\boldsymbol{s}_k))^2}, \quad \text{MAE}_X = \frac{1}{53}\sum_{k=1}^{53}|Z_X(\boldsymbol{s}_k) - \widehat{Z}_X(\boldsymbol{s}_k)|,$$

where $\widehat{Z}_X(\boldsymbol{s}_k)$ is the cokriging predictor of the variable $X = T, P$ (temperature and precipitation respectively) at the point on the sphere $\boldsymbol{s}_k$, $k = 1, \ldots, 53$ based on 200×2 observations. We repeat 500 times and record all RMSE's and MAE's for each variable. Table 3 reports the empirical mean of the five hundred RMSE's and MAE's for the different bivariate covariance models, for each variable. We denote it with $\overline{\text{RMSE}}_X$ and $\overline{\text{MAE}}_X$, $X = T, P$. We can see that the $\mathcal{F}$ model outperforms the chordal Matérn and the the Generalized Wendland models for both measures of prediction performance in the separable and nonseparable cases.

Model	l_N	ρ_{12}	σ_{11}^2	σ_{22}^2	α			ν			$\overline{\text{RMSE}}_T$	$\overline{\text{RMSE}}_P$	$\overline{\text{MAE}}_T$	$\overline{\text{MAE}}_P$
$\mathcal{BW}^S$	-708.96	0.3165	2.5710	12.214	0.5476	-	-	1.0527	-	-	0.5325	1.1485	0.5321	1.1474
$\mathcal{BM}^S$	-704.98	0.3231	2.5049	11.744	0.0507	-	-	2.2058	-	-	0.5289	1.1373	0.5285	1.1362
$\mathcal{BF}^S$	-704.21	0.3231	2.5269	11.820	0.0673	-	-	2.4426	-	-	0.5281	1.1341	0.5276	1.1329
Model	l_N	ρ_{12}	σ_{11}^2	σ_{22}^2	α_{11}	α_{22}	α_{12}	ν_{11}	ν_{22}	ν_{12}	$\overline{\text{RMSE}}_T$	$\overline{\text{RMSE}}_P$	$\overline{\text{MAE}}_T$	$\overline{\text{MAE}}_P$
$\mathcal{BW}^{NS}$	-699.36	0.2836	2.3518	12.352	0.5567	0.5677	1.4603	0.9279	0.9728	0.1442	0.5211	1.1310	0.5207	1.1299
$\mathcal{BM}^{NS}$	-697.24	0.2897	2.3822	12.416	0.0599	0.0628	0.2128	1.8029	1.7950	0.7096	0.5191	1.1198	0.5187	1.1187
$\mathcal{BF}^{NS}$	-697.04	0.2788	2.3436	11.442	0.0733	0.0750	0.2488	2.1410	2.1378	0.7925	0.5181	1.1186	0.5177	1.1175

Table 3: Loglikelihood (l_N, see Equation (12)) and ML estimates for separable and nonseparable Wendland, Matérn and $\mathcal{F}$ bivariate models with associated RMSE and MAE for the the first and second variable.

5 Discussion

This paper has limited its scope to bivariate random fields. Under the multivariate parametric adaptation construction principle, the trivariate case would already imply severe restrictions on the colocated correlation coefficients ρ_{ij}, for $i, j = 1, 2, 3$ $(i \neq j)$. This is a well known problem in multivariate spatial modeling and we refer to (Gneiting et al., 2010) and to Daley et al. (2015) for constructive criticism. A way to circumvent this problem might be to adapt convolution based approaches that have proved to be successful (Gneiting et al., 2010). We are not aware of any approach for multivariate covariances based on convolution on the sphere. For scalar valued fields,

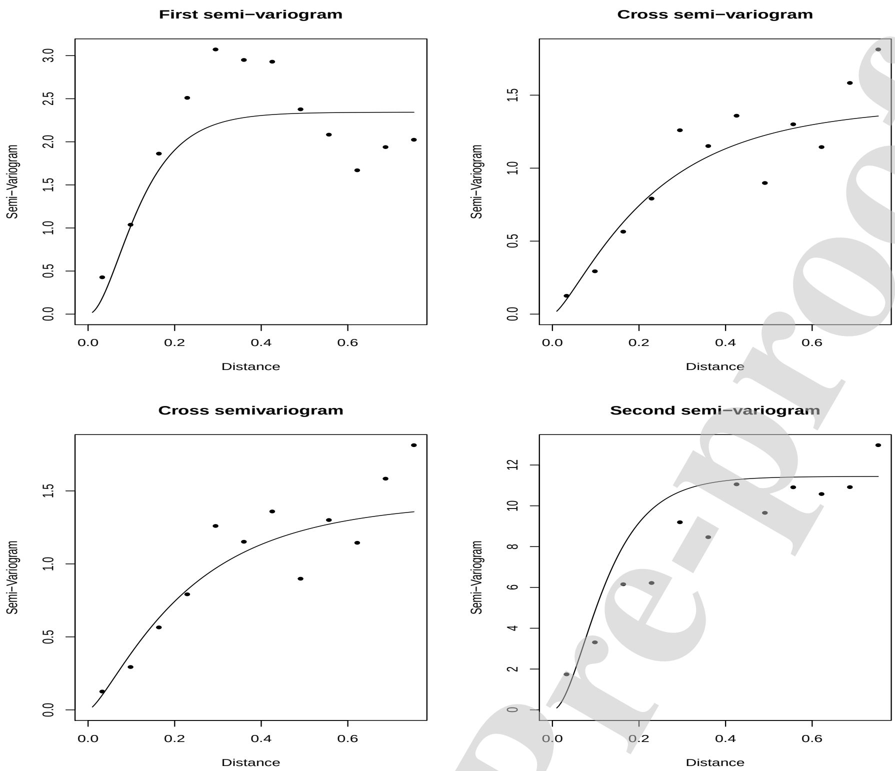


Figure 5: Dots indicate empirical marginal semi-variograms (diagonal) and cross-variograms (off diagonal) estimations for the temperature and precipitation data compared to the estimated semi-variograms under the nonseparable bivariate $\mathcal{F}$ model.

the work of Hansen et al. (2015) has a promising approach to convolution based covariance functions on spheres. Adapting this approach might lead to a good solution. In the simulation study and in the application we adopt the ML estimation method. This kind of estimation can be computationally expensive, in particular when working with data on the whole Earth surface. In this case, other types of estimation methods for multivariate random fields, having a good balance between statistical efficiency and computational complexity, might be considered. For instance, composite likelihood (Bevilacqua et al., 2016a), covariance tapering (Furrer et al., 2016; Bevilacqua et al., 2016b) or the methods outlined in Heaton et al. (2019) might be considered.

Acknowledgement

Partial support was provided by FONDECYT grant 1130647, Chile for Emilio Porcu. Partial support was provided by FONDECYT grant 1200068, Chile for Moreno Bevilacqua and by Millennium Science Initiative of the Ministry of Economy, Development, and Tourism, grant "Millenium Nucleus Center for the Discovery of Structures in Complex Data" for Moreno Bevilacqua and Emilio Porcu. We thank the Referees and the Associate Editor for the thorough work that allowed for a clearly improved version of the manuscript.

A Appendix: Proof of Theorem 2.1

Proof. We start by proving **(A)**, by direct construction. We start by reporting Lemma 1 in Daley et al. (2015) and rephrased for our convenience.

Lemma A.1 *Let* $\boldsymbol{G}(\cdot) : (0,1) \to \mathbb{R}^{2\times 2}$ *be the matrix valued mapping having elements*

$$g_{ij}(\delta) = \frac{\rho_{ij}}{B(\alpha_{ij}, \nu_{ij})} \delta^{\alpha_{ij}-1}(1-\delta)^{\nu_{ij}-1}, \qquad \delta \in (0,1), \quad i,j = 1,2, \tag{13}$$

where $\rho_{ii} = 1$ *and* $\rho_{12} = \rho_{21}$. *If* $\alpha_{12} = \frac{1}{2}(\alpha_{11} + \alpha_{22})$, *then* $\boldsymbol{G}(\delta)$ *is positive definite for any fixed* $\delta \in (0,1)$ *if and only if*

- $\nu_{11} + \nu_{22} - 2\nu_{12} \leq 0$;
- $|\rho_{12}| \leq \frac{B(\alpha_{12},\nu_{12})}{(B(\alpha_{11},\nu_{11})B(\alpha_{22},\nu_{22}))^{1/2}}$.

For comparison with Daley et al. (2015), our Equation (13) corresponds to Equation 4 and 5 in Daley et al. (2015). Also, the constants b_{ij} in Lemma 1 of Daley et al. (2015) are identically equal to one in our case; the constant c_{ij} in Lemma 1 of Daley et al. (2015) are equal to $\rho_{ij}/B(\alpha_{ij}, \nu_{ij})$. Finally,

the constants γ_{ij} in Lemma 1 of Daley et al. (2015) correspond to $\nu_{ij} - 1$ in our case.
Some notation is now needed to illustrate the proof. Let Ψ_d^2 be the class of mappings $\boldsymbol{C} : [0, \pi] \to \mathbb{R}^{2\times 2}$ having elements C_{ij} that are pointwise continuous, such that $\boldsymbol{C}(\theta)$ is the geodesically isotropic covariance function of a bivariate Gaussian random field $\boldsymbol{Z}$ defined on the d-dimensional unit sphere $\mathbb{S}^d$ embedded in $\mathbb{R}^{d+1}$. Call $\Psi_\infty^2 := \bigcap_{d\geq 1} \Psi_d^2$. Let Ψ_d be the class of pointwise continuous mappings $C : [0, \pi] \to \mathbb{R}$ such that $C(\theta)$ is a geodesically isotropic covariance function of a scalar random field in $\mathbb{S}^d$. Accordingly, we define $\Psi_\infty := \bigcap_{d\geq 1} \Psi_d$.
The classes Ψ_d and Ψ_d^2 are nested, with the strict inclusion relations

$$\Psi_1 \supset \Psi_2 \supset \ldots \supset \Psi_\infty \quad \text{and} \quad \Psi_1^2 \supset \Psi_2^2 \supset \ldots \supset \Psi_\infty^2.$$

Arguments in Gneiting (2013) show that $C \in \Psi_\infty$ if and only if

$$C(\theta) = \sum_{k=0}^{\infty} b_k (\cos\theta)^k, \qquad \theta \in [0, \pi],$$

where the coefficients b_k are nonnegative and summable. This implies, for instance, that the mapping $\mathcal{N}_{\delta,\tau}$ defined at (9) belongs to the class Ψ_∞ for any $\tau > 0$ and $\delta \in (0, 1)$, because

$$\mathcal{N}_{\delta,\tau}(\theta) = \sum_{k=0}^{\infty} b_k(\delta, \tau)(\cos\theta)^k, \qquad \theta \in [0, \pi],$$

with

$$b_k(\delta, \tau) = \binom{k + \tau - 1}{k} \delta^k (1 - \delta)^\tau, \qquad \delta \in (0, 1), \tau > 0,$$

and $k = 0, 1, \ldots$. See, for instance, Equation 16 in Gneiting (2013) or Theorem 6.4 in DasGupta (2010). The class Ψ_∞^2 has instead been characterized in Hannan (2009) and Yaglom (1987) through the expansion

$$\boldsymbol{C}(\theta) = \sum_{k=0}^{\infty} \boldsymbol{B}_k (\cos\theta)^k, \qquad \theta \in [0, \pi], \tag{14}$$

where $\{\boldsymbol{B}_k\}_{k=0}^{\infty}$ is an absolutely convergent sequence of positive definite matrices (summability is intended pointwise).

Another technical result is now needed.

Lemma A.2 *Let $\delta \in (0,1)$, $\tau > 0$. Let $\mathcal{N}_{\delta,\tau} : [0,\pi] \to \mathbb{R}$ be defined as in (9). Let $\widetilde{\boldsymbol{\lambda}} = (\delta,\tau)^\top$ and consider vectors $\widetilde{\boldsymbol{\lambda}}_{ij} = (\delta_{ij},\tau_{ij})^\top$, $i,j = 1,2$ with $\delta_{ij} \in (0,1)$, $\tau_{ij} > 0$ and such that $\widetilde{\boldsymbol{\lambda}}_{12} = \widetilde{\boldsymbol{\lambda}}_{21}$. Let $\widetilde{\boldsymbol{C}} : [0,\pi] \to \mathbb{R}^{2\times 2}$ with elements $\widetilde{C}_{ij}$ being defined through*

$$\widetilde{C}_{ij}(\theta) = \widetilde{\rho}_{ij}\mathcal{N}_{\delta_{ij},\tau_{ij}}(\theta), \qquad \theta \in [0,\pi],$$

where $\widetilde{\rho}_{ii} = 1$ and $\widetilde{\rho}_{12} = \widetilde{\rho}_{21}$. Then, $\widetilde{\boldsymbol{C}} \in \Psi^2_\infty$ provided $\delta_{12} \le \min\{\delta_{11},\delta_{22}\}$, $\tau_{12} \le \min\{\tau_{11},\tau_{22}\}$, and

$$|\widetilde{\rho}_{12}| \le \frac{(1-\delta_{11})^{\tau_{11}/2}(1-\delta_{22})^{\tau_{22}/2}}{(1-\delta_{12})^{\tau_{12}}}.$$

The proof of Lemma A.2 comes straight from identity (9) together with characterization (14). In fact, we need to show that the matrices $\boldsymbol{B}_k$ with elements $b_{i,j,k} = \widetilde{\rho}_{ij} b_k(\delta_{ij},\tau_{ij})$, $i,j = 1,2$, $k = 0,1,\ldots$, are positive definite for all k, and form an absolutely convergent sequence (Hannan, 2009). This is in turn ensured by solving the determinantal inequality

$$\widetilde{\rho}_{12}^2 \le \frac{(1-\delta_{11})^{\tau_{11}}(1-\delta_{22})^{\tau_{22}}}{(1-\delta_{12})^{2\tau_{12}}} \inf_{k=0,1,\ldots} \frac{(\tau_{11}+k-1)_k(\tau_{22}+k-1)_k}{(\tau_{12}+k-1)_k^2}\left(\frac{\delta_{11}\delta_{22}}{\delta_{12}^2}\right)^k,$$

where $(x)_k = x(x+1)\cdots(x+k-1)$, $x \in \mathbb{R}$ (Abramowitz and Stegun, 1970). The infimum, with respect to $k = 0,1,\ldots$, on the right hand side is attained at $k = 0$ when $\tau_{12} \le \min(\tau_{11},\tau_{22})$ and $\delta_{12} \le \min(\delta_{11},\delta_{22})$. To show it, we have

$$\begin{aligned}\frac{(\tau_{11}+k-1)_k(\tau_{22}+k-1)_k}{(\tau_{12}+k-1)_k^2} &= \frac{\tau_{11}(\tau_{11}+1)\cdots(\tau_{11}+k-1)}{\tau_{12}(\tau_{12}+1)\cdots(\tau_{12}+k-1)}\frac{\tau_{22}(\tau_{22}+1)\cdots(\tau_{22}+k-1)}{\tau_{12}(\tau_{12}+1)\cdots(\tau_{12}+k-1)} \\ &= \prod_{n=0}^{k-1}\left(\frac{\tau_{11}-\tau_{12}}{\tau_{12}+n}+1\right)\prod_{n=0}^{k-1}\left(\frac{\tau_{22}-\tau_{12}}{\tau_{12}+n}+1\right),\end{aligned}$$

which is strictly increasing in k. Thus, the infimum on the right hand side is attained at $k = 0$, where it is identically equal to one. This completes the proof of the lemma.

We now provide a criterion that is the crux of the proof. A proof is not provided, as it can be obtained by applying the same arguments as in Porcu and Zastavnyi (2011).

Lemma A.3 *Let $\mathcal{N}_{\delta,\tau} : [0,\pi] \to \mathbb{R}$ be defined as in Equation (9). Let $\boldsymbol{G}$ be the matrix valued function with elements g_{ij} defined through (13). If the constants $\tau_{ij}, \alpha_{ij}, \nu_{ij}, \rho_{ij}$, $i,j = 1,2$, satisfy the requirements of Lemmas A.1 and A.2, then the mapping*

$$\theta \mapsto \sigma_{ii}\sigma_{jj}\rho_{ij} \int_0^1 \mathcal{N}_{\delta,\tau_{ij}}(\theta)\boldsymbol{G}(\delta)\mathrm{d}\delta, \tag{15}$$

belongs to the class Ψ_∞^2.

Arguments in the proof of Theorem 3.1 in Alegría et al. (2018) show that

$$\mathcal{F}_{\tau_{ij},\alpha_{ij},\nu_{ij}}(\theta) = \int_{(0,1)} \mathcal{N}_{\delta,\tau_{ij}}(\theta) g_{ij}(\delta)\mathrm{d}\delta, \qquad \theta \in [0,\pi], \tag{16}$$

where $g_{ij}(\cdot)$ has been defined at Lemma A.1. We now note that Lemma A.2 under $\delta_{ij} = \delta \in (0,1)$ implies $\widetilde{\rho}_{12} \le 1$, so that we can pick $\widetilde{\rho}_{12} = 1$ and fix it throughout. The proof is thus completed by coupling (16) with Lemmas A.1, A.2 and A.3.

To prove assertion (**B**), we invoke again the characterization in Equation (14), and resort to the scale mixture (16) in concert with Equation 16 in Alegría et al. (2018) to find that the mapping $\boldsymbol{C}$ in Equation (8) can be uniquely written as in Equation (14), with

$$b_{ij,k} = b_k(\boldsymbol{\lambda}_{ij}) = \sigma_{ii}\sigma_{jj}\rho_{ij} \frac{B(\alpha_{ij}, \nu_{ij}+\tau_{ij})}{B(\alpha_{ij},\nu_{ij})} \frac{(\alpha_{ij})_k(\tau_{ij})_k}{(\alpha_{ij}+\nu_{ij}+\tau_{ij})_k k!}, \qquad k = 0,1,\ldots.$$

Showing positive definiteness of $\boldsymbol{B}_k$ for all k amounts to solving the determinantal inequality

$$\rho_{12}^2 \le \frac{B(\alpha_{11},\nu_{11}+\tau_{11})B(\alpha_{22},\nu_{22}+\tau_{22})}{B(\alpha_{12},\nu_{12}+\tau_{12})^2} \frac{B(\alpha_{12},\nu_{12})^2}{B(\alpha_{11},\nu_{11})B(\alpha_{11},\nu_{22})}\Upsilon, \tag{17}$$

with

$$\Upsilon = \inf_{k=0,1,\ldots} \frac{(\alpha_{11})_k(\alpha_{22})_k}{(\alpha_{12})_k^2} \frac{(\tau_{11})_k(\tau_{22})_k}{(\tau_{12})_k^2} \frac{(\alpha_{12}+\nu_{12}+\tau_{12})_k^2}{(\alpha_{11}+\nu_{11}+\tau_{11})_k(\alpha_{22}+\nu_{22}+\tau_{22})_k}. \tag{18}$$

We now prove that under Conditions 1 and 2 in Assertion **(B)**, such a infimum is uniquely attained at $k=0$. In fact, direct inspection shows that

$$\begin{aligned}
\frac{(\alpha_{11})_k(\alpha_{22})_k}{(\alpha_{12})_k^2} &= \frac{\alpha_{11}(\alpha_{11}+1)\cdots(\alpha_{11}+k-1)}{\alpha_{12}(\alpha_{12}+1)\cdots(\alpha_{12}+k-1)} \frac{\alpha_{22}(\alpha_{22}+1)\cdots(\alpha_{22}+k-1)}{\alpha_{12}(\alpha_{12}+1)\cdots(\alpha_{12}+k-1)} \\
&= \prod_{n=0}^{k-1}\left(\frac{\alpha_{11}-\alpha_{12}}{\alpha_{12}+n}+1\right)\prod_{n=0}^{k-1}\left(\frac{\alpha_{22}-\alpha_{12}}{\alpha_{12}+n}+1\right),
\end{aligned}$$

which is strictly increasing in k provided Condition 1 in Assertion **(B)** holds. Analogously, one can show that the second and third factors in the right hand side of Equation (18) are strictly increasing in k provided that Condition $\tau_{12} \leq min\{\tau_{11}, \tau_{22}\}$ and Condition 2 in Assertion (B) hold, respectively. □

References

Abramowitz, M. and Stegun, I. A., editors (1970). *Handbook of Mathematical Functions*. Dover, New York.

Alegría, A., Caro, S., Bevilacqua, M., Porcu, E., and Clarke, J. (2017). Estimating covariance functions of multivariate Skew-Gaussian random fields on the sphere. *Spatial Statistics*, 22:388 – 402.

Alegría, A., Cuevas, F., Diggle, P., and Porcu, E. (2018). A new class of covariance functions of random fields on spheres. *Technical Report, University of Aalborg. Available at https://data.math.au.dk/publications/csgb/2018/math−csgb−2018−08.pdf.*

Alegría, A. and Porcu, E. (2017). The dimple problem related to space–time modeling under the Lagrangian framework. *Journal of Multivariate Analysis*, 162:110 – 121.

Alegría, A., Porcu, E., Furrer, R., and Mateu, J. (2019). Covariance functions for multivariate Gaussian fields. *Stochastic Environmental Research Risk Assessment*, 33:1593–1608.

Apanasovich, T. V. and Genton, M. G. (2010). Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika*, 97:15 –30.

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(497):180–193.

Banerjee, S. (2005). On geodetic distance computations in spatial modeling. *Biometrics*, 61(2):617–625.

Bevilacqua, M., Alegria, A., Velandia, D., and Porcu, E. (2016a). Composite likelihood inference for multivariate Gaussian random fields. *Journal of Agricultural, Biological, and Environmental Statistics*, 21(3):448–469.

Bevilacqua, M., Faouzi, T., Furrer, R., and Porcu, E. (2019). Estimation and prediction using Generalized Wendland covariance functions under fixed domain asymptotics. *Annals of Statistics*, 47(2):828–856.

Bevilacqua, M., Fassó, A., Gaetan, C., Porcu, E., and Velandia, D. (2016b). Covariance tapering for multivariate gaussian random fields estimation. *Statistical Methods and Applications*, 25(1):21–37.

Bevilacqua, M. and Morales-Oñate, V. (2018). *GeoModels: A Package for Geostatistical Gaussian and non Gaussian Data Analysis*. R package version 1.0.3-4.

Bevilacqua, M., Vallejos, R., and Velandia, D. (2015). Assessing the significance of the correlation between the components of a bivariate gaussian random field. *Environmetrics*, 26(8):545–556.

Combes, V., Hormázabal, S., and Di Lorenzo, E. (2017). Interannual variability of the subsurface eddy field in the southeast pacific. *Journal of Geophysical Research-Oceans*, 120:2769–2783.

Daley, D., Porcu, E., and Bevilacqua, M. (2015). Classes of compactly supported covariance functions for multivariate random fields. *Stochastic Environmental Research and Risk Assessment*, 29(4):1249–1263.

DasGupta, A. (2010). *Fundamentals of Probability: A First Course*. Springer, Berlin.

Di Lorenzo, E., Combes, V., Keister, J., Strub, P., Andrew, T., Peter, F., Marck, O., Furtado, J. ., Bracco, A., Bograd, S., Peterson, W., Schwing, F., Taguchi, B., Hormázabal, S., and Parada, C. (2014). Synthesis of pacific ocean climate and ecosystem dynamics. *Oceanography*, 26(4):68–81.

Edwards, M., Castruccio, S., and Hammerling, D. (2019). A multivariate global spatio-temporal stochastic generator for climate ensembles. *Journal of Agricultural, Biological and Environmental Sciences*.

Furrer, R., Bachoc, F., and Du, J. (2016). Asymptotic properties of multivariate tapering for estimation and prediction. *J. Multivar. Anal.*, 149(C):177–191.

Furrer, R., Genton, M. G., and Nychka, D. (2006). Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics*, 15:502–523.

Genton, M. G. and Kleiber, W. (2015). Cross-covariance functions for multivariate geostatistics. *Statistical Science*, 30(2):147–163.

Gneiting, T. (2013). Strictly and non-strictly positive definite functions on spheres. *Bernoulli*, 19(4):1327–1349.

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.

Hannan, E. (2009). *Multiple Time Series*. Wiley Series in Probability and Statistics. Wiley.

Hansen, L. V., Thorarinsdottir, T. L., Ovcharov, E., Gneiting, T., and Richards, D. (2015). Gaussian random particles with flexible Hausdorff dimension. *Advances in Applied Probability*, 47(2):307–327.

Heaton, M. J., Datta, A., Finley, A. O., Furrer, R., Guinness, J., Guhaniyogi, R., Gerber, F., Gramacy, R. B., Hammerling, D., Katzfuss, M., Lindgren, F., Nychka, D. W., Sun, F., and

Zammit-Mangion, A. (2019). A case study competition among methods for analyzing large spatial data. *J. Agric. Biol. Environ. Stat.*, (in press):1–28.

Li, B. and Zhang, H. (2011). An approach to modeling asymmetric multivariate spatial covariance structures. *Journal of Multivariate Analysis*, 102(10):1445 – 1453.

Nychka, D., Bandyopadhyay, S., Hammerling, D., Lindgren, F., and Sain, S. (2015). A multi-resolution Gaussian process model for the analysis of large spatial datasets. *Journal of Computational and Graphical Statistics*, 24(2):579–599.

Porcu, E., Alegría, A., and Furrer, R. (2018). Modeling spatially global and temporally evolving data. *International Statistical Review*, 86:344–377.

Porcu, E., Bevilacqua, M., and Genton, M. G. (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.

Porcu, E., Gregori, P., and Mateu, J. (2006). Nonseparable stationary anisotropic space–time covariance functions. *Stochastic Environmental Research and Risk Assessment*, 21(2):113–122.

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(9):1293 – 1301.

Reinsel, G., Tiao, G., Wang, M., Lewis, R., and Nychka, D. (1981). Statistical analysis of stratospheric ozone data for the detection of trends. *Atmospheric Environment (1967)*, 15(9):1569–1577.

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*, 73:992–103.

Wackernagel, H. (2003). *Multivariate Geostatistics: An Introduction with Applications.* Springer, New York, 3rd edition.

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

Yaglom, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions. Volume I: Basic Results.* Springer, New York.

Zhang, H. and Wang, Y. (2010). Kriging and cross-validation for massive spatial data. *Environmetrics*, 21(3-4):290–304.