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

Online Supplement

Emilio Porcu,[1] Moreno Bevilacqua,[2] and Marc G. Genton[3]

June 4, 2015

Abstract

In this paper, 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 in order 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.

Key words: Chordal distance; Cokriging; Global data; Great circle distance; Multivariate; Spatio-temporal statistics; Sphere.

Short title: Covariance Functions of the Great Circle Distance

[1]Deparment of Mathematics, University Federico Santa Maria, 2360102 Valparaiso, Chile. E-mail: emilio.porcu@usm.cl
Research supported by Proyecto Fondecyt Regular from Chilean Government of Education.
[2]Universidad de Valparaiso, Department of Statistics, 2360102 Valparaiso, Chile.
E-mail: moreno.bevilacqua@uv.cl

[3]CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia.
E-mail: marc.genton@kaust.edu.sa

Tables of Completely Monotone and Bernstein Functions

Table 1: Possible choices for the function $0 < t \mapsto \varphi(t)$ being completely monotone on the positive real line. The range of parameters allows for complete monotonicity. $\mathcal{K}_\nu$ denotes a modified Bessel function. Note that all of the mappings can be rescaled with a positive constant

Family	Analytic Expression	Parameter Range
Powered Exponential	$\varphi(t) = \exp\{-t^\gamma\}$	$\gamma \in (0,1]$
Matérn	$\varphi(t) = \frac{2^{\nu-1}}{\Gamma(\nu)} t^\nu \mathcal{K}_\nu t^\nu$	$\nu \in (0, 1/2]$
Generalized Cauchy	$\varphi(t) = \{1+t^\gamma\}^{-\nu}$	$\gamma \in (0,1]; \quad \nu > 0$
Dagum	$\varphi(t) = 1 - \left\{\frac{t^{\tau\delta}}{(c+t^\tau)^\delta}\right\}$	$\tau \in (0,1]; \quad \delta \in (0,1]$

Table 2: Examples of Bernstein functions. The function $\Gamma(a;t) := \int_t^\infty t^{a-1}\,\mathrm{e}^{-t}\,dt$ is the incomplete Gamma function. All the examples are taken from Porcu and Schilling (2011).

Function	Parameters Restriction	Function	Parameters Restriction
$(1+t^\alpha)^\beta$	$0 < \alpha, \beta \leq 1$	$\mathrm{e}t - t\Big(1+\frac{1}{t}\Big)^t - \frac{t}{t+1}$	
$\Big(\frac{t^\rho}{1+t^\rho}\Big)^\gamma$	$0 < \gamma, \rho < 1$	$\frac{1}{a} - \frac{1}{t}\log\Big(1+\frac{t}{a}\Big)$	$a > 0$
$\frac{t^\alpha - t(1+t)^{\alpha-1}}{(1+t)^\alpha - t^\alpha}$	$0 < \alpha < 1$	$\sqrt{\frac{t}{2}}\,\frac{\sinh^2\sqrt{2t}}{\sinh(2\sqrt{2t})}$	
$\sqrt{t}\left(1-\mathrm{e}^{-2a\sqrt{t}}\right)$	$a > 0$	$t^{1-\nu}\,\mathrm{e}^{at}\,\Gamma(\nu; at)$	$a > 0,\ 0 < \nu < 1$
$\frac{t\left(1-\mathrm{e}^{-2\sqrt{t+a}}\right)}{\sqrt{t+a}}$	$a > 0$	$t^\nu\,\mathrm{e}^{a/t}\,\Gamma\big(\nu; \frac{a}{t}\big)$	$a > 0,\ 0 < \nu < 1$

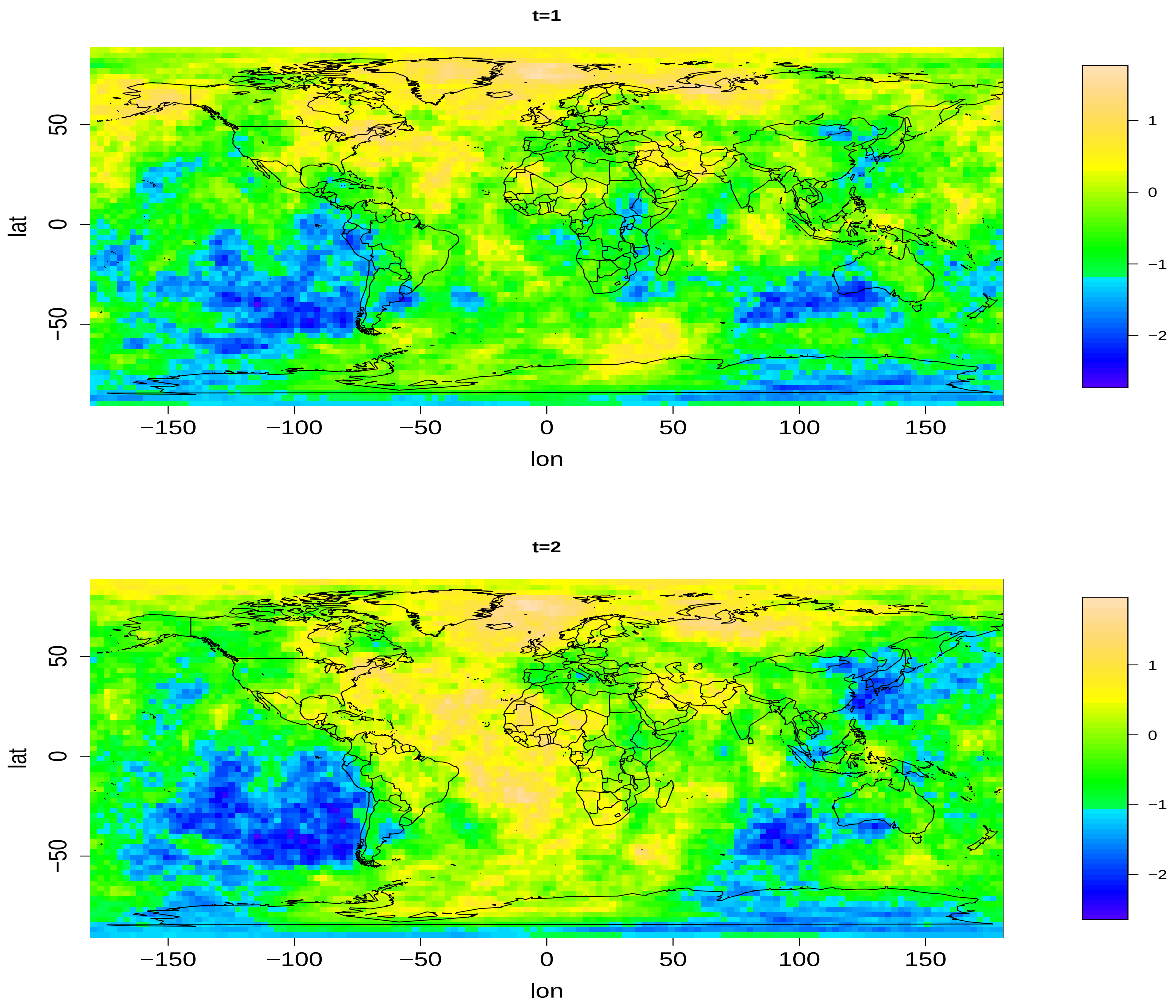

Figure 1: Realization of a space-time Gaussian process (at two different time instants) with the covariance in (15) in the published manuscript (PM throughout).

Space-Time Covariance from Direct Construction

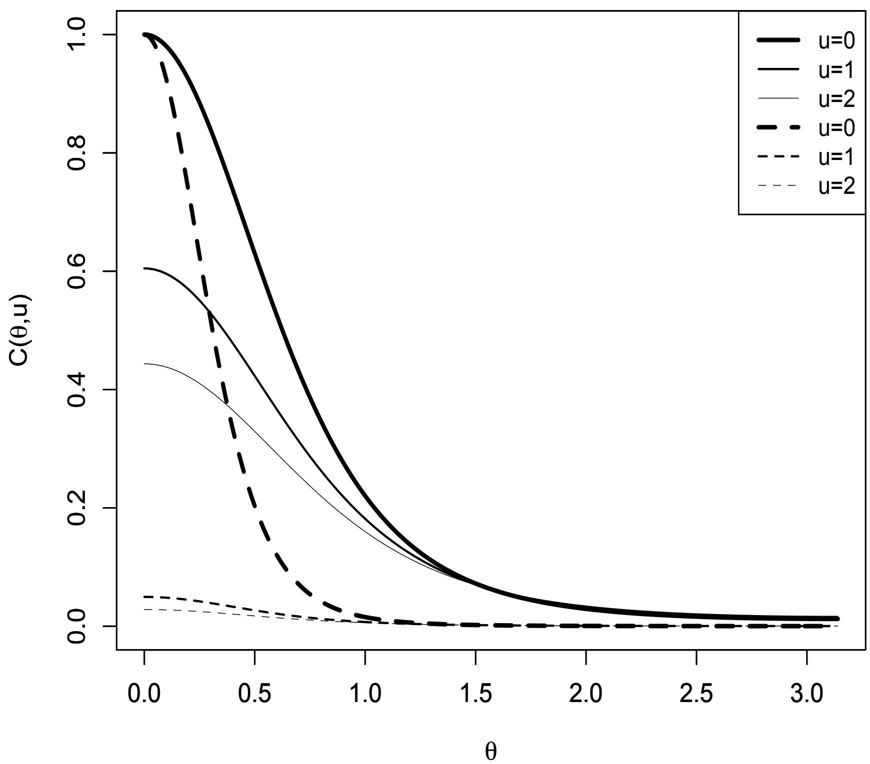

Figure 2: Plots of the covariances $C(\theta, u)$ from (11) (solid) and from (12) in the PM (dashed).

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

Prediction

We now assess possible discrepancies between GC, CH, and MP distances in terms of prediction. We propose an illustrative example obtained by working in the purely spatial case. In particular, we consider three scenarios (called, respectively, (A), (B) and (C)) with a decreasing number of location sites, with $N = 81, 25$, and 4 (see Figure 3).

For each scenario, we simulate, using GC distance, 10,000 spatial Gaussian random fields with zero mean, unit variance, and with exponential covariance:

$$C(\theta; a) = \exp(-R\theta/a), \quad (1)$$

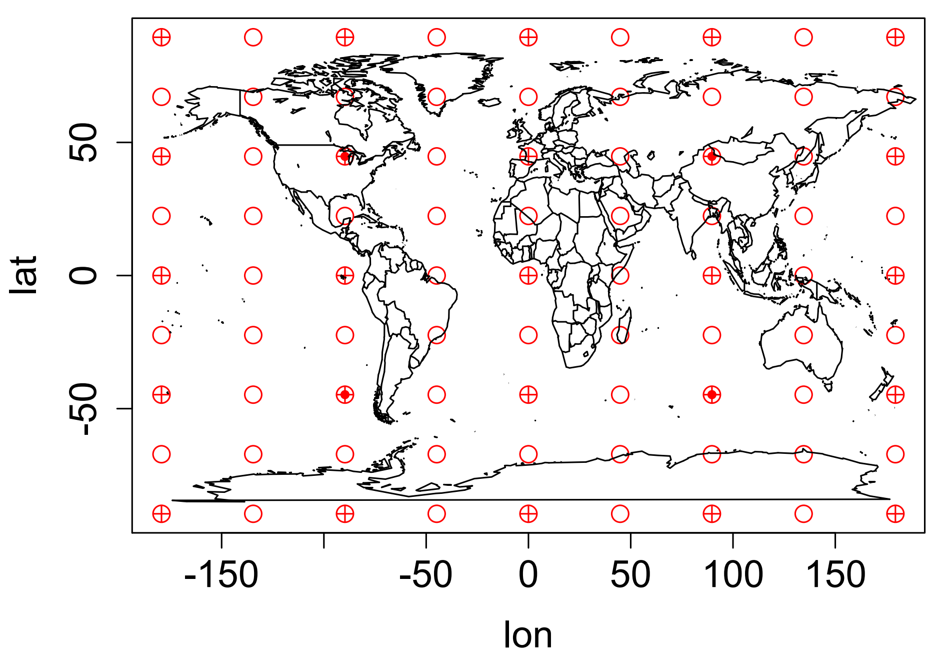

Figure 3: Scenarios considered in the simulation study to evaluate the predictive scores: (A) 81 open circles ∘; (B) 25 plus signs +; (C) 4 closed circles •.

where we set the parameter a equal to $10,000/3$ or $20,000/3$. Thus, we have six cases (two for each scenario), each with observations $Z^{(k)}(\mathbf{x}_i)$ for $i = 1, \ldots, N$, where the index $k = 1, \ldots, 10,000$ denotes the replicates of the simulation.

We then consider for each scenario the quantity

$$\hat{Z}_a^{(k)}(\mathbf{x}_i, \mathcal{X}), \quad \text{where } \mathcal{X} = \mathrm{GC}, \mathrm{CH}, \mathrm{MP}, \qquad k = 1, \ldots, 10,000, \tag{2}$$

that is the best linear prediction of $Z^{(k)}(\mathbf{x}_i)$ (that depends on the parameter a) based on all data $Z^{(k)}(\mathbf{x}_l)$ for $l \neq i$ using the GC, CH or MP distance. Since the goal is to assess how the choice of the distance affects the prediction, we consider the covariance parameters as known. We propose three predictive scores for each simulation (Zhang and Wang, 2010; Gneiting and Raftery, 2007) based on (2). The first prediction score is the root mean square error (RMSE):

$$\mathrm{RMSE}^{(k)}(a, \mathcal{X}) = \left[\frac{1}{N}\sum_{i=1}^{N}\left\{Z^{(k)}(\mathbf{x}_i) - \hat{Z}_a^{(k)}(\mathbf{x}_i, \mathcal{X})\right\}^2\right]^{1/2}, \qquad \mathcal{X} = \mathrm{GC}, \mathrm{CH}, \mathrm{MP}. \tag{3}$$

The second prediction score is the logarithmic score:

$$\mathrm{log}S^{(k)}(a, \mathcal{X}) = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{2}\mathrm{log}\{2\pi\sigma^{(k)}(\mathbf{x}_i)\} + \frac{1}{2}\{Y_a^{(k)}(\mathbf{x}_i)\}^2\right], \qquad \mathcal{X} = \mathrm{GC}, \mathrm{CH}, \mathrm{MP}, \tag{4}$$

where $Y_a^{(k)}(\mathbf{x}_i) = \frac{Z^{(k)}(\mathbf{x}_i) - \hat{Z}_a^{(k)}(\mathbf{x}_i, \mathcal{X})}{\sigma^{(k)}(\mathbf{x}_i)}$ and $\{\sigma^{(k)}(\mathbf{x}_i)\}^2$ is the prediction variance associated with $\hat{Z}_a{}^{(k)}(\mathbf{x}_i)$. Finally, we consider the continuous rank probability score (CRPS), which can be written in the Gaussian case as

$$\mathrm{CRPS}^{(k)}(a, \mathcal{X}) = \frac{1}{N}\sum_{i=1}^{N}\sigma^{(k)}(\mathbf{x}_i)\left(Y_a^{(k)}(\mathbf{x}_i)\left[2\Phi\left\{Y_a^{(k)}(\mathbf{x}_i)\right\} - 1\right] + 2\Phi\left\{Y_a^{(k)}(\mathbf{x}_i)\right\} - \frac{1}{\sqrt{\pi}}\right), \tag{5}$$

for $\mathcal{X} = \mathrm{GC}, \mathrm{CH}, \mathrm{MP}$, where Φ is the Gaussian cumulative distribution function.

In Table 3, we report the overall means (summing up with respect to $k = 1, \ldots, 10,000$) of the prediction score rules above for scenarios (A), (B), and (C) and $a = 10,000/3,\ 20,000/3$. The prediction scores tend to show similar behavior when the number of points is increased when using GC and CH distances while they become worse when using the MP distance. The

Table 3: Averaged RMSE$(a, \mathcal{X})$, logS$(a, \mathcal{X})$ and CRPS$(a, \mathcal{X})$ for $\mathcal{X} = \mathrm{GC}, \mathrm{CH}, \mathrm{MP}$, $a = 10,000/3, 20,000/3$ for scenarios (A), (B) and (C).

Scenario	$(\frac{10,000}{3}, \mathrm{GC})$	$(\frac{10,000}{3}, \mathrm{CH})$	$(\frac{10,000}{3}, \mathrm{MP})$	$(\frac{20,000}{3}, \mathrm{GC})$	$(\frac{20,000}{3}, \mathrm{CH})$	$(\frac{20,000}{3}, \mathrm{MP})$
(A) RMSE	0.593	0.593	0.705	0.437	0.437	0.537
(B) RMSE	0.576	0.577	0.727	0.468	0.470	0.617
(C) RMSE	0.806	0.818	0.799	0.803	0.824	0.797
Scenario	$(\frac{10,000}{3}, \mathrm{GC})$	$(\frac{10,000}{3}, \mathrm{CH})$	$(\frac{10,000}{3}, \mathrm{MP})$	$(\frac{20,000}{3}, \mathrm{GC})$	$(\frac{20,000}{3}, \mathrm{CH})$	$(\frac{20,000}{3}, \mathrm{MP})$
(A) logS	0.438	0.438	0.781	0.117	0.118	0.493
(B) logS	0.061	0.062	0.721	-0.229	-0.228	0.492
(C) logS	1.303	1.314	1.296	1.292	1.312	1.288
Scenario	$(\frac{10,000}{3}, \mathrm{GC})$	$(\frac{10,000}{3}, \mathrm{CH})$	$(\frac{10,000}{3}, \mathrm{MP})$	$(\frac{20,000}{3}, \mathrm{GC})$	$(\frac{20,000}{3}, \mathrm{CH})$	$(\frac{20,000}{3}, \mathrm{MP})$
(A) CRPS	1.099	1.099	1.364	0.807	0.808	1.025
(B) CRPS	0.951	0.950	1.386	0.758	0.759	1.168
(C) CRPS	2.022	2.030	2.017	1.959	1.962	1.977

scenario (C), with four isolated points, shows the biggest difference between GC, CH, and MP distances. As expected, the strength of the spatial dependence affects the prediction using different distances. Specifically, with stronger dependence, GC performs better than CH and MP in terms of prediction, albeit the differences are very small in the CH case.

TOMS Data: Prediction

We also assess the prediction performance of the covariance models proposed in our estimation analysis through the predictive scores defined in Equations (3), (4) and (5) of the OS. The results are presented in Table 4.

Table 4: RMSE, logS, and CRPS using GC, CH and MP distances for TOMS data using the covariance models A.1, A.2, B.1, B.2, B.3, and C.1 in the PM.

A.1		A.1		A.1	
RMSE($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	6.4905	logS($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	3.2617	CRPS($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	12.5516
RMSE($\hat{\boldsymbol{\lambda}}_{GC}, CH$)	6.4907	logS($\hat{\boldsymbol{\lambda}}_{GC}, CH$)	3.2618	CRPS($\hat{\boldsymbol{\lambda}}_{GC}, CH$)	12.5511
RMSE($\hat{\boldsymbol{\lambda}}_{GC}, MP$)	6.7374	logS($\hat{\boldsymbol{\lambda}}_{GC}, MP$)	3.2864	CRPS$\hat{\boldsymbol{\lambda}}_{GC}, MP$)	12.9262
RMSE($\hat{\boldsymbol{\lambda}}_{CH}, GC$)	6.4906	logS($\hat{\boldsymbol{\lambda}}_{CH}, GC$)	3.2617	CRPS($\hat{\boldsymbol{\lambda}}_{CH}, GC$)	12.5399
RMSE($\hat{\boldsymbol{\lambda}}_{CH}, CH$)	6.4908	logS($\hat{\boldsymbol{\lambda}}_{CH}, CH$)	3.2618	CRPS($\hat{\boldsymbol{\lambda}}_{CH}, CH$)	12.5394
RMSE($\hat{\boldsymbol{\lambda}}_{CH}, MP$)	6.7376	logS($\hat{\boldsymbol{\lambda}}_{CH}, MP$)	3.2864	CRPS($\hat{\boldsymbol{\lambda}}_{CH}, MP$)	12.9144
RMSE($\hat{\boldsymbol{\lambda}}_{MP}, GC$)	6.5318	logS($\hat{\boldsymbol{\lambda}}_{MP}, GC$)	3.2664	CRPS($\hat{\boldsymbol{\lambda}}_{MP}, GC$)	12.3273
RMSE($\hat{\boldsymbol{\lambda}}_{MP}, CH$)	6.5319	logS($\hat{\boldsymbol{\lambda}}_{MP}, CH$)	3.2665	CRPS($\hat{\boldsymbol{\lambda}}_{MP}, CH$)	12.3268
RMSE($\hat{\boldsymbol{\lambda}}_{MP}, MP$)	6.7700	logS($\hat{\boldsymbol{\lambda}}_{MP}, MP$)	3.2849	CRPS($\hat{\boldsymbol{\lambda}}_{MP}, MP$)	12.7562
A.2		A.2		A.2	
RMSE($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	6.5295	logS($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	3.2647	CRPS($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	12.6020
RMSE($\hat{\boldsymbol{\lambda}}_{GC}, CH$)	6.5297	logS($\hat{\boldsymbol{\lambda}}_{GC}, CH$)	3.2648	CRPS($\hat{\boldsymbol{\lambda}}_{GC}, CH$)	12.6014
RMSE($\hat{\boldsymbol{\lambda}}_{GC}, MP$)	6.7758	logS($\hat{\boldsymbol{\lambda}}_{GC}, MP$)	3.2869	CRPS($\hat{\boldsymbol{\lambda}}_{GC}, MP$)	13.0298
RMSE($\hat{\boldsymbol{\lambda}}_{CH}, GC$)	6.5329	logS($\hat{\boldsymbol{\lambda}}_{CH}, GC$)	3.2654	CRPS($\hat{\boldsymbol{\lambda}}_{CH}, GC$)	12.5352
RMSE($\hat{\boldsymbol{\lambda}}_{CH}, CH$)	6.5331	logS($\hat{\boldsymbol{\lambda}}_{CH}, CH$)	3.2655	CRPS($\hat{\boldsymbol{\lambda}}_{CH}, CH$)	12.5346
RMSE($\hat{\boldsymbol{\lambda}}_{CH}, MP$)	6.7778	logS($\hat{\boldsymbol{\lambda}}_{CH}, MP$)	3.2869	CRPS($\hat{\boldsymbol{\lambda}}_{CH}, MP$)	12.9663
RMSE($\hat{\boldsymbol{\lambda}}_{MP}, GC$)	6.5860	logS($\hat{\boldsymbol{\lambda}}_{MP}, GC$)	3.2719	CRPS($\hat{\boldsymbol{\lambda}}_{MP}, GC$)	12.3168
RMSE($\hat{\boldsymbol{\lambda}}_{MP}, CH$)	6.5861	logS($\hat{\boldsymbol{\lambda}}_{MP}, CH$)	3.2719	CRPS($\hat{\boldsymbol{\lambda}}_{MP}, CH$)	12.3163
RMSE($\hat{\boldsymbol{\lambda}}_{MP}, MP$)	6.8243	logS($\hat{\boldsymbol{\lambda}}_{MP}, MP$)	3.2861	CRPS($\hat{\boldsymbol{\lambda}}_{MP}, MP$)	12.7997
B.1		B.1		B.1	
RMSE($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	6.4314	logS($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	3.2518	CRPS($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	12.4359
B.2		B.2		B.2	
RMSE($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	6.4301	logS($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	3.2516	CRPS($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	12.4382
B.3		B.3		B.3	
RMSE($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	6.3953	logS($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	3.2477	CRPS($\hat{\boldsymbol{\lambda}}_{GC}, GC$)	12.3860
C.1		C.1		C.1	
RMSE($\hat{\boldsymbol{\lambda}}_{CH}, CH$)	6.3191	logS($\hat{\boldsymbol{\lambda}}_{CH}, CH$)	3.2407	CRPS($\hat{\boldsymbol{\lambda}}_{CH}, CH$)	12.3586
RMSE($\hat{\boldsymbol{\lambda}}_{CH}, MP$)	6.5510	logS($\hat{\boldsymbol{\lambda}}_{CH}, MP$)	3.2673	CRPS($\hat{\boldsymbol{\lambda}}_{CH}, MP$)	12.6598
RMSE($\hat{\boldsymbol{\lambda}}_{MP}, CH$)	6.3689	logS($\hat{\boldsymbol{\lambda}}_{MP}, CH$)	3.2446	CRPS($\hat{\boldsymbol{\lambda}}_{MP}, CH$)	12.1704
RMSE($\hat{\boldsymbol{\lambda}}_{MP}, MP$)	6.5885	logS($\hat{\boldsymbol{\lambda}}_{MP}, MP$)	3.2636	CRPS($\hat{\boldsymbol{\lambda}}_{MP}, MP$)	12.5628

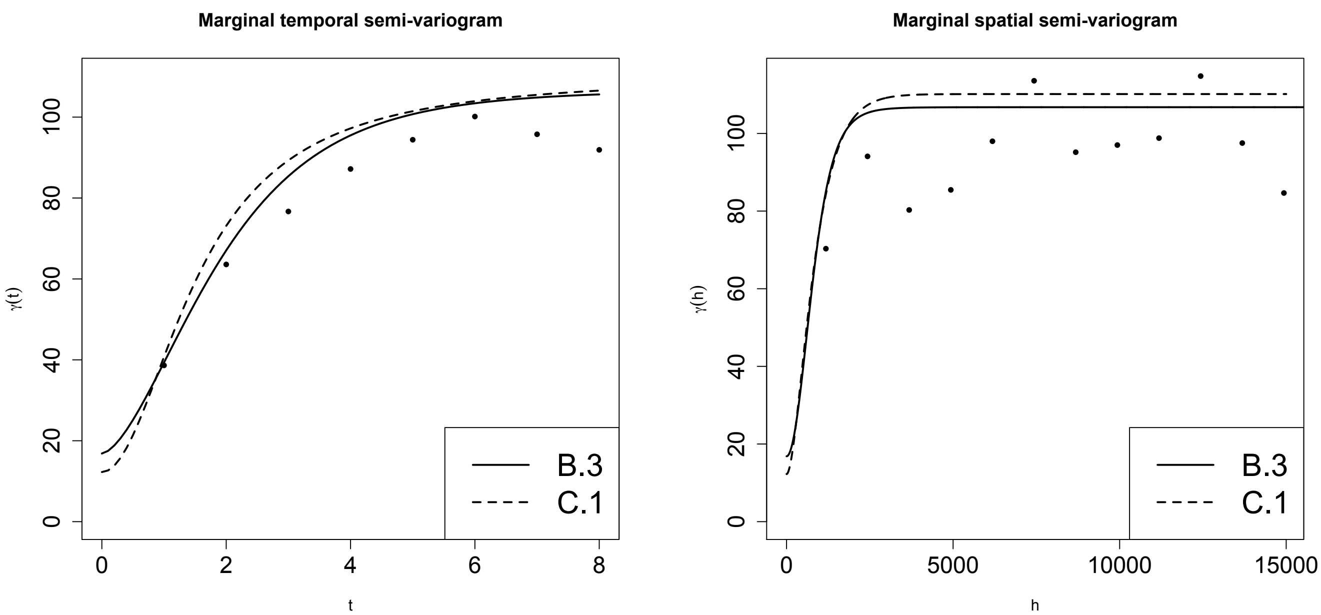

Figure 4: Marginal empirical temporal semivariogram vs. estimated temporal semivariogram (left) and marginal empirical spatial semivariogram vs. estimated spatial semivariogram (right) using model B.3 and C.1 in the PM.

Appendix A: the class $\Psi_{d,T}$

A1. A Gneiting class on the sphere cross time

We start with a technical lemma that we use in the proof of Theorem 1.

Lemma 1 *Let $(\Omega, \mathcal{A}, \mu)$ be a measure space, with μ a nonnegative measure. Let $f(\cdot;\cdot) : [0,\pi] \times \Omega \to \mathbb{R}$ and $g(\cdot;\cdot) : \mathbb{R} \times \Omega \to \mathbb{R}$ such that*

1. *The product $f(\theta;\cdot) \times g(u;\cdot)$ belongs to $L_1(\Omega, \mathcal{A}, \mu)$ a.e. every $\theta \in [0,\pi]$ and $u \in \mathbb{R}$.*

2. *For d a positive integer, and every $\tau_0 \in \Omega$, $f(\cdot;\tau)$ belongs to the class Ψ_d, and $g(\cdot;\tau)$ is a continuous positive definite function on the real line.*

Then the function $K : [0,\pi] \times \mathbb{R} \to \mathbb{R}$ defined by

$$K(\theta, u) = \int_{\Omega} f(\theta;\tau) g(u;\tau) \mu(\mathrm{d}\tau), \qquad \theta \in [0,\pi], \quad u \in \mathbb{R}, \tag{6}$$

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

Proof. By Assumption 2, the function $K_\tau(\cdot,\cdot) := f(\cdot;\tau)g(\cdot;\tau)$ belongs to the class $\Psi_{d,T}$ a.e. $\tau \in \Omega$. Assumption 1 implies that the function K is well defined. In order to show continuity of K, note that $K_\tau(\cdot,\cdot)$ is continuous and that, a.e. $\tau \in \Omega$, we have that $\lim_{(\theta,u)\to(0,0)} K_\tau(\theta,u) = K_\tau(0,0)$. Also we have that $|K_\tau(\theta,u)| \le K_\tau(0,0)$. By Lebesgue's theorem, we get continuity of K. The proof is then completed by noting that positive definite functions are closed under scale mixtures, since the set of positive definite functions is a convex cone closed under the topology of positive and finite measures. $\square$

The Gneiting class on the great circle: proof of Theorem 1

Consider the mapping

$$(u,\tau) \mapsto g_a(u;\tau) := \cos(|u|\tau) \exp\left(-a\tau^2\right), \qquad a > 0, \qquad (u,\tau) \in \mathbb{R} \times [0,\infty).$$

Let the measure μ in Lemma 1 be the Lebesgue measure on the positive real line. Clearly, Assumption 2 in Lemma 1 is satisfied by Schoenberg's representation of symmetric positive definite functions on the real line (Schoenberg, 1938). Now, for ν a positive and bounded measure on the positive real line, we consider the function

$$(\theta,\tau) \mapsto f(\theta;\tau) := \frac{1}{\sqrt{\pi}} \int_{[0,\infty)} \exp\left\{-\frac{\tau^2 \psi_{[0,\pi]}(\theta)}{4r}\right\} \nu(\mathrm{d}r), \tag{7}$$

where $\psi_{[0,\pi]}$ is the restriction to the interval $[0,\pi]$ of a positive-valued Bernstein function, and ν is a positive and bounded measure. Application of Fubini's theorem shows that Assumption 1 in Lemma 1 is satisfied. Regarding Assumption 2, consider the mapping $h_r(\theta;\tau) := \exp\left\{-\frac{\tau^2}{4r}\psi_{[0,\pi]}(\theta)\right\}$. By a criterion in Feller (Feller, 1966, p. 441), the composition of the negative exponential function with a Bernstein function is a completely monotone function on the positive real line, for any positive τ, r. Direct application of Theorem 7 in Gneiting (2013) shows that $h_r(\cdot;\tau)$ is the restriction of a completely monotone function to the interval $[0,\pi]$ and thus belongs to the class Ψ_∞ a.e. $\tau, r > 0$. Thus, the function $f(\cdot;\cdot)$ in Equation (7) satisfies Assumption 2 in Lemma 1. We are now able to apply Lemma 1 and we have

$$\begin{aligned}
C_a(\theta,u) &:= \int_0^\infty g_a(u;\tau) f(\theta;\tau)\mathrm{d}\tau \\
&= \int_0^\infty \cos(|u|\tau)\exp\left(-a\tau^2\right)\left[\frac{1}{\sqrt{\pi}}\int_{[0,\infty)} \exp\left\{-\frac{\tau^2\psi_{[0,\pi]}(\theta)}{4r}\right\}\nu(\mathrm{d}r)\right]\mathrm{d}\tau \\
&= \frac{1}{\sqrt{\pi}}\int_{[0,\infty)}\int_0^\infty \cos(|u|\tau)\exp\left[-\tau^2\left\{a+\frac{\psi_{[0,\pi]}(\theta)}{4r}\right\}\right]\mathrm{d}\tau\nu(\mathrm{d}r) \\
&= \int_{[0,\infty)}\left\{\frac{r}{\psi_{[0,\pi]}(\theta)+4ar}\right\}^{1/2}\exp\left\{-\frac{ru^2}{\psi_{[0,\pi]}(\theta)+4ar}\right\}\nu(\mathrm{d}r). \qquad (8)
\end{aligned}$$

We have that $C_a \in \Psi_{\infty,T}$ for any $a > 0$. Since positive definite functions are a convex cone closed under the topology of pointwise convergence, we have, using dominated convergence, that

$$C(\theta,u) = \lim_{a\to 0} C_a(\theta,u) = \int_{[0,\infty)}\left\{\frac{1}{\psi_{[0,\pi]}(\theta)}\right\}^{1/2}\exp\left\{-\frac{ru^2}{\psi_{[0,\pi]}(\theta)}\right\}\widetilde{\nu}(\mathrm{d}r),$$

with $\widetilde{\nu}(\mathrm{d}r) := \sqrt{r}\nu(\mathrm{d}r)$. Invoking Bernstein's theorem (Feller, 1966, p.441), we can write C as

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

for φ completely monotone on the positive real line. The proof is completed. $\square$

The modified Gneiting class on the great circle: proof of Theorem 2

We provide a proof of the constructive type, based on the following arguments: we follow Gneiting (2013) when defining

$$\varphi_{c,n}(\theta) = \left(1 - \frac{\theta}{c}\right)_+^{n+1}, \qquad \theta \in [0, \pi],$$

where $c \in (0, 2\pi]$, $n \leq 3$ is an integer, and where $(\cdot)_+$ denotes the positive part of a real argument. By Lemmas 3 and 4 in Gneiting (2013), $\varphi_{c,n} \in \Psi_{2n+1}$ when $c \in (0, \pi]$ and $\varphi_{c,n} \in \Psi_\infty$ when $c > \pi$, for any integer $n \geq 0$. Thus, $\varphi_{c,n}$ belongs to the class Ψ_{2n+1} for $n \leq 3$ and for all $c > 0$. We define the mapping $u \mapsto g(u; \xi, n, m) := \xi^{n+1}\{1 - \xi\psi(|u|)\}_+^m$, $u \in \mathbb{R}$. A criterion of the Pólya type (Gneiting, 2001) shows that g is positive definite on the real line for any $\xi \in \mathbb{R}_+$, for all integers n, m. The mappings $\varphi_{c,n}$ and g as defined satisfy the requirements of Lemma 1, so that we can again invoke the mixture argument in Equation (6) and have that the mapping

$$C_{n,m}(\theta, u; c) = \int_0^\infty \varphi_{\xi c,n}(\theta) g(u; \xi, n, m) \mathrm{d}\xi, \qquad (\theta, u) \in [0, \pi] \times \mathbb{R},$$

belongs to the class $\Psi_{2n+1,T}$, $n \leq 3$. Let $B(\cdot, \cdot)$ denote the Beta function. The arguments in Porcu and Zastavnyi (2014) and direct inspection show that

$$C_{n,m}(\theta, u; c) = B(n+2, m+1) \frac{1}{\psi(|u|)^{n+2}} \varphi_{c,n+m+1} \left\{\theta\psi(|u|)\right\}, \qquad (\theta, u) \in [0, \pi] \times \mathbb{R},$$

and for the following we omit the factor $B(n+2, m+1)$ because it does not affect positive definiteness. Positive definite functions are a convex cone, closed under the topology of pointwise convergence. Thus, we have that

$$C_n(\theta, u; c) = \lim_{m \to \infty} C_{n,m}(\theta, u, m/c) = \frac{1}{\psi(|u|)^{n+2}} \exp\left\{-c\theta\psi(|u|)\right\}, \qquad (\theta, u) \in [0, \pi],$$

belongs to the class $\Psi_{2n+1,T}$, for all $a > 0$. Let $\mu(\cdot)$ be a positive and bounded measure. Again by scale mixture argument, we have that

$$\widetilde{C}_n(\theta, u) = \frac{1}{\psi(|u|)^{n+2}} \int_0^\infty \exp\{-a\theta\psi(|u|)\}\, \mu(\mathrm{d}a), \qquad (\theta, u) \in [0, \pi] \times \mathbb{R},$$

is a member of the class $\Psi_{2n+1,T}$. Invoking again Bernstein's theorem (Feller, 1966, p.441):

$$\widetilde{C}_n(\theta, u) = \frac{1}{\psi(|u|)^{n+2}} \varphi\{\theta\psi(|u|)\}, \qquad (\theta, u) \in [0, \pi] \times \mathbb{R},$$

where φ is a completely monotone. This proves the assertion. $\square$

A2. Direct constructions on the sphere cross time

We present a more general result than the one exposed in Section 2.2 of the PM for the class $\Psi_{d,T}$. For the following, let $\mathbf{x} \in \mathbb{S}^d \subset \mathbb{R}^{d+1}$. Let $Y_{k,\nu,d}(\cdot) : \mathbb{S}^d \to \mathbb{C}$, $k = 0, 1, \ldots$, $\nu \in \Upsilon_{k,d}$, define the normalized hyperspherical harmonics (for a standard reference, see Dai and Xu (2013), Ch. 1). Here and throughout, $\Upsilon_{k,d}$ is an index set such that $Y_{k,\nu,d}$, $k = 0, \ldots$, form an orthonormal basis of $L_2(\mathbb{S}^d, \nu_d)$ with ν_d being the surface area measure. For given $\lambda > 0$ and k a nonnegative integer, we also define the Gegenbauer polynomials P_k^λ through (Dai and Xu, 2013)

$$\frac{1}{(1 + r^2 - 2r\cos\theta)^\lambda} = \sum_{k=0}^\infty r^k P_k^\lambda(\cos\theta), \qquad \theta \in [0, \pi], \quad r \in (-1, 1).$$

By Schoenberg's Theorem (Schoenberg, 1942), a mapping $C : [0, \pi] \to \mathbb{R}$ is an element of the class Ψ_d if and only if

$$C(\theta) = \sum_{k=0}^\infty b_{k,d} \frac{P_k^{(d-1)/2}(\cos\theta)}{P_k^{(d-1)/2}(1)}, \qquad \theta \in [0, \pi],$$

where we use the terminology in Daley and Porcu (2013) and call the probability mass sequence $\{b_{k,d}\}_{k=0}^\infty$ a d-Schoenberg sequence. Analogously, $C \in \Psi_\infty$ if and only if

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

where the Schoenberg sequence $\{b_k\}_{k=0}^{\infty}$ does not depend on d.

Theorem 1 *Let $\{g_k(\cdot)\}_{k=0}^{\infty}$ be an absolutely summable sequence of continuous positive definite functions on the real line, with $g_k(0) = b_{k,d}$, $k = 0, 1, \ldots$, with $\{b_{k,d}\}_{k=0}^{\infty}$ a d-Schoenberg sequence. Then the series expansion*

$$C(\theta, u) = \sum_{k=0}^{\infty} g_k(u) \frac{P_k^{(d-1)/2}(\cos\theta)}{P_k^{(d-1)/2}(1)}, \qquad (\theta, u) \in [0, \pi] \times \mathbb{R}, \tag{9}$$

defines a representation for elements of the class $\Psi_{d,T}$. Accordingly,

$$C(\theta, u) = \sum_{k=0}^{\infty} g_k(u)(\cos\theta)^k, \qquad (\theta, u) \in [0, \pi] \times \mathbb{R}, \tag{10}$$

defines a representation for members of the class $\Psi_{\infty,T}$, where in this case $g_k(0) = b_k$ for a Schoenberg sequence $\{b_k\}_{k=0}^{\infty}$.

Proof. We define a Gaussian process Z through the series expansion

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

where we assume that the set of all $\xi_{k,\nu}(t)$, $t \in \mathbb{R}$, forms a countable infinite sequence of Gaussian processes, with zero mean and $\mathbb{E}\xi_{k,\nu}(t)\xi_{k',\nu'}(t') = \delta_{k,k'}\delta_{\nu,\nu'} g_k(t - t')$, $t, t' \in \mathbb{R}$, where $\{g_k(\cdot)\}_{k=0}^{\infty}$ is an absolutely summable sequence of continuous positive definite functions defined on the real line with $g_k(0) = b_{k,d}$ for a d-Schoenberg sequence $b_{k,d}$. Here $\delta_{k,k'}$ denotes the Kronecker delta.

This series representation is well defined in view of completeness of the normalized hyperspherical harmonics. Observe that the sequence of Gaussian processes $\xi_{k,\nu}$ determines the finite dimensional distributions of the process Z defined on $\mathbb{S}^d \times \mathbb{R}$. We can now invoke the addition theorem for hyperspherical harmonics (Dai and Xu, 2013), in concert with their orthonormality property, so that

$$P_k^{(d-1)/2}(\cos\theta) = c_d \sum_{\nu \in \Upsilon_{k,d}} Y_{k,\nu,d}(\mathbf{x}) \overline{Y_{k,\nu,d}(\mathbf{x}')}, \qquad \mathbf{x}, \mathbf{x}' \in \mathbb{S}^d,$$

with c_d a positive constant depending on dimension d. Equation (9) is then verified through direct inspection. A convergence argument as much as in Schoenberg (1942) then shows Equation (10). The proof is completed. $\square$

A3. Verifying the examples in Table 1 (PM)

Negative Binomial According to Newton's generalized binomial theorem, for $a \in (0,1)$, and for $\tau > 0$, the mapping $x \mapsto (1-ax)^{-\tau}$ has positive Taylor coefficients. In particular:

$$(1-ax)^{-\tau} = \sum_{k=0}^{\infty} \binom{\tau+k-1}{k} a^k x^k,$$

with $|ax| < 1$ being the necessary condition for the absolute convergence of the series, and

$$\binom{\tau+k-1}{k} = \frac{(\tau+k-1)_k}{k!},$$

and where $(\cdot)_k$ is the Pochhammer symbol. Thus, the negative binomial class can be written

$$\sum_{k=0}^{\infty} \binom{\tau+k-1}{k} \varepsilon^k g(u)^k (\cos\theta)^k.$$

Multiquadric The proof comes trivially by setting $c := 2\varepsilon/(1+\varepsilon^2) \in (0,1)$ in the negative binomial expansion.

Sine Series The mapping $t \mapsto \exp(t)(1+t)$ admits a series expansion of the type

$$\sum_{k=0}^{\infty} \frac{(k+1)}{k!} t^k, \qquad t \in \mathbb{R}.$$

The rest comes from the same arguments as in previous points.

Sine Power We make use of the following constructive, arguments. First, again by the generalized binomial theorem and extending the arguments in Soubeyrand et al. (2008) we have that, for $\alpha \in (0,2]$,

$$\frac{1}{2^{\alpha/2}}\{1-g(u)\cos\theta\}^{\alpha/2} = \frac{1}{2^{\alpha/2}}\left\{1-\sum_{k=1}^{\infty} \tau_k(\alpha/2) g(u)^k (\cos\theta)^k\right\},$$

where $\tau_k(x) = -\frac{1}{k!}\prod_{m=0}^{k-1}(m-x)$. As shown in Soubeyrand et al. (2008), the coefficients are strictly positive whenever $\alpha \in (0,2)$. This completes the proof. Observe that, by the well known identity $\sin(\theta/2) = \sqrt{(1-\cos\theta)/2}$, we obtain that the spatial margin $C(\theta,0)$ is exactly the Sine Power family as introduced in Soubeyrand et al. (2008).

Adapted Multiquadric Given the assumption on the function $h(\cdot) := 2g(\cdot)/\{1+g^2(\cdot)\}$, we have, by the arguments in the proof of the negative binomial family, that the mapping

$$C(\theta,u) = \left\{\frac{1}{1-\varepsilon h(u)\cos\theta}\right\}^{\tau}, \qquad (\theta,u) \in [0,\pi]\times\mathbb{R},$$

belongs to the class $\Psi_{\infty,T}$, for $\varepsilon \in (0,1)$ and for any positive τ. Plugging in $h(\cdot) = 2g(\cdot)/\{1+g^2(\cdot)\}$ we obtain the result.

Appendix B: the class Ψ_∞^p

Diagonal dominance conditions for the Matérn and Cauchy classes

This section is devoted to extend the results in Section 3 to the case $p > 2$. We focus on the example of the Matérn and Cauchy covariances, being more popular in spatial statistics. The following results are based on the same technique exposed through Theorem A: for a given completely monotone function φ, we consider its restriction $\varphi_{[0,\pi]}$ and then adopt the parameterization as detailed through Equation (13) in the PM. Thus, for the mapping $\boldsymbol{C}$ in Equation (14) in the PM to be in the class Φ_∞^p, we need the associated measure $\boldsymbol{\mu}(\xi_0)$ to belong to Ψ_0^p for each ξ_0 in $[0,\infty)$. The measures associated to the Matérn and Generalized Cauchy functions (actually, their Laplace transforms) are reported in Table 2 (PM) (first and second entries, respectively). For the cases $p > 2$, a simple determinantal inequality is not sufficient for showing positive definiteness. A more restrictive assumption of diagonal dominance is needed to show our results.

Matérn cross-covariance models on the sphere

For a self-contained exposition, we report here the expression of the Matérn function, being completely monotonic only for $\nu \in (0, 1/2]$. We use a slightly different parameterization:

$$\mathcal{M}(t;\alpha,\nu,\rho) = \rho\frac{2^{1-\nu}}{\Gamma(\nu)}(\alpha t)^\nu \mathcal{K}_\nu(\alpha t), \qquad t \geq 0, \tag{11}$$

where $\alpha > 0$ is a scale parameter, ν is a positive parameter indexing smoothness, and $\mathcal{K}_\nu$ a modified Bessel function. We denote by $\mathcal{M}_{[0,\pi]}$ the restriction of $\mathcal{M}$ to the interval $[0,\pi]$.

The use of the parameter ρ will become clearer from the exposition of the subsequent results. In Equation (11), ρ is strictly positive but its range of validity can be modified according to subsequent results. For the case of Euclidean distances, the mapping $\boldsymbol{C}$ with Matérn members and with parameterization as in Equation (14) in the PM has been analyzed in Gneiting et al. (2010) and Apanasovich et al. (2012). Unfortunately, their results are not useful for the construction in this paper. As mentioned, we make use of Theorem A, in concert with diagonal dominance. For the following, we write $\mu(\mathrm{d}\xi) = g(\xi)\mathrm{d}\xi$ for the Matérn measure as in Table 2 (PM) (first entry, fifth column), so that we can write $g(\cdot; \alpha, \nu, \rho)$ for the mapping

$$(0, \infty) \ni \xi \mapsto g(\xi; \alpha, \nu, \rho) := \rho \left(\frac{\alpha^2}{4}\right)^{\nu} \frac{\xi^{-1-\nu}}{\Gamma(\nu)} \exp\left(-\frac{\alpha^2}{4\xi}\right). \tag{12}$$

Denote $[g_{ij}(\xi) := g(\xi; \alpha_{ij}, \nu_{ij}, \rho_{ij})]_{i,j=1}^{p}$ by $\mathbf{g}(\xi)$, and set $\rho_{ii} = 1$ for obvious reasons. The arguments in Theorem A show that, if $\mathbf{g}(\xi) \in \Psi_0^p$ for every fixed value of $\xi \in (0, \infty)$, the matrix-valued mapping

$$[0, \pi] \ni \theta \mapsto \mathbf{C}(\theta) = \left[\mathcal{M}_{[0,\pi]}(t; \alpha_{ij}, \nu_{ij}, \rho_{ij}) = \int_{[0,\infty)} \exp(-\theta^2 \xi) g_{ij}(\xi)\mathrm{d}\xi\right]_{i,j=1}^{p} \tag{13}$$

belongs to the class Ψ_d^p for any positive integers d and p.

Two technical lemmas are useful for the main result of this section.

Lemma 2 *Let $\rho > 0$ and $\rho_1 \in \mathbb{R}$. For p a positive integer, with p larger than 2, a necessary and sufficient condition for*

$$g(\xi; \alpha, \nu, \rho) \geq |g(\xi; \alpha_1, \nu_1, \rho_1)|(p-1), \qquad \xi > 0, \tag{14}$$

is that $\nu_1 \geq \nu > 0$, $\alpha_1 \geq \alpha > 0$, and

$$|\rho_1| \leq \frac{\rho}{(p-1)} \frac{\Gamma(\nu_1)}{\Gamma(\nu)} \frac{\alpha^{2\nu}}{\alpha_1^{2\nu_1}} \left\{\exp(1) \frac{\alpha_1^2 - \alpha^2}{\nu_1 - \nu}\right\}^{\nu_1 - \nu}.$$

Proof. We borrow the idea from the technique in Theorem 4 of Gneiting et al. (2010). The assertion follows by noting that $g(\cdot; \alpha, \nu, \rho) \geq |g(\cdot; \alpha_1, \nu_1, \rho_1)|(p-1)$ on $(0, \infty)$ if and only if

$$|\rho_1| \leq \frac{\rho}{(p-1)} \frac{\Gamma(\nu_1)}{\Gamma(\nu)} \frac{\alpha^{2\nu}}{\alpha_1^{2\nu_1}} \inf \left\{(4\xi)^{\nu_1 - \nu} \exp\left(\frac{\alpha_1^2 - \alpha^2}{4\xi}\right)\right\}.$$

The rest of the proof comes from the same argument in Gneiting et al. (2010). □

Lemma 3 *The matrix $[g_{ij}(\xi)]_{i,j=1}^{p}$, with g_{ij}'s defined via (12), is diagonally dominant for any fixed $\xi > 0$, if $\nu_{ji} = \nu_{ij} \geq \nu_{ii} > 0$, $\alpha_{ji} = \alpha_{ij} \geq \alpha_{ii} > 0$, $\rho_{ij} = \rho_{ji}$ $(j \neq i)$, $\rho_{ii} = 1$, and*

$$|\rho_{ij}| \le \frac{1}{(p-1)} \frac{\Gamma(\nu_{ij})}{\Gamma(\nu_{ii})} \frac{\alpha_{ii}^{2\nu_{ii}}}{\alpha_{ij}^{2\nu_{ij}}} \left\{ \exp(1) \frac{\alpha_{ii}^2 - \alpha_{ij}^2}{\nu_{ij} - \nu_{ii}} \right\}^{\nu_{ij}-\nu_{ii}}, \qquad j \neq i, \tag{15}$$

with the limiting case

$$|\rho_{ij}| \le \frac{1}{(p-1)} \frac{\alpha_{ii}^{2\nu_{ii}}}{\alpha_{ij}^{2\nu_{ii}}}, \qquad j \neq i,$$

obtained when $\nu_{ij} = \nu_{ii}$. In particular, diagonal dominance implies that such a matrix belongs to the class Ψ_0^p, for any positive integer $p \ge 2$.

Proof. From the conditions in (15), sum for $j \neq i$ both sides of the inequality to find that $g_{ii}(\xi) \ge \sum_{j \neq i} |g_{ij}(\xi)|$ for all $\xi \ge 0$. □

Theorem 2 *Let $0 < \nu_{ij} \le 1/2$, $i, j = 1, \ldots, p$. Under the conditions in Lemmas 2 and 3, the matrix-valued mapping defined by*

$$\widetilde{\mathbf{C}}(\theta) := \left[\widetilde{C}_{ij}(\theta)\right]_{i,j=1}^{p} = \left[\sigma_i \sigma_j \rho_{ij} C_{ij}(\theta)\right]_{i,j=1}^{p}, \tag{16}$$

with $\mathbf{C}(\theta)$ defined at (13), belongs to the class Ψ_d^p for any positive integers d and $p \ge 2$. The σ_i^2's define the variances of the processes Z_i, $i = 1, \ldots, p$ and the ρ_{ij}'s are colocated correlation coefficients.

Cauchy cross-covariance models on the sphere

The previous techniques can be used to generate matrix-valued covariances that can exhibit heavy or light tails, i.e., covariances that respectively have long or short range dependence. Two more technical lemmas are useful in this section. We first define:

$$[0, \infty) \ni \xi \mapsto h(\xi; \nu, \mu, \rho) := \rho \frac{\mu^{\nu}}{\Gamma(\nu)} \xi^{\nu-1} \exp(-\mu\xi), \qquad \nu, \mu > 0, \quad \rho \in \mathbb{R}, \tag{17}$$

and denote $\boldsymbol{h}(\xi) := [h_{ij}(\xi) = h(\xi; \nu_{ij}, \mu_{ij}, \rho_{ij})]_{i,j=1}^{p}$, $\xi \ge 0$, where again $\rho_{ii} = 1$. Observe that h is the density, with respect to the Lebesgue measure, of the generalized Cauchy function as detailed through the second entry of Table 2 of the PM.

Lemma 4 *Let $\rho > 0$ and $\rho_1 \in \mathbb{R}$. A necessary and sufficient condition for*

$$h\,(\xi;\nu,\mu,\rho) \geq (p-1)|h\,(\xi;\nu_1,\mu_1,\rho_1)\,|, \qquad \xi > 0, \tag{18}$$

is that $\nu_1 \geq \nu > 0$, $\mu_1 \geq \mu > 0$, and

$$|\rho_1| \leq \frac{\rho}{(p-1)}\frac{\Gamma(\nu_1)}{\Gamma(\nu)}\frac{\mu^{\nu}}{\mu_1^{\nu_1}}\left\{\exp(-1)\frac{\nu_1-\nu}{\mu_1-\mu}\right\}^{\nu-\nu_1}.$$

The proof is omitted since it is obtained along the same lines as for Lemma 2.

Lemma 5 *If $\nu_{ji} = \nu_{ij} \geq \nu_{ii}$, $\mu_{ji} = \mu_{ij} \geq \mu_{ii}$, $\rho_{ij} = \rho_{ij}$ $(j \neq i)$, $\rho_{ii} = 1$, and*

$$|\rho_{ij}| \leq \frac{1}{(p-1)}\frac{\Gamma(\nu_{ij})}{\Gamma(\nu_{ii})}\frac{\mu_{ii}^{\nu_{ii}}}{\mu_{ij}^{\nu_{ij}}}\left\{\exp(-1)\frac{\nu_{ij}-\nu_{ii}}{\mu_{ij}-\mu_{ii}}\right\}^{\nu_{ii}-\nu_{ij}},$$

then the matrix $[h_{ij}(\xi)]_{i,j=1}^{p}$, with $h_{ij}(\cdot)$ defined through Equation (17), is diagonally dominant for each $\xi \geq 0$. Thus, it belongs to the class Ψ_0^p for any positive integer $p \geq 2$.

We now have all the ingredients to offer a matrix-valued mapping of the generalized Cauchy type (Gneiting and Schlather, 2004). We define

$$\mathcal{C}(t;\nu,\mu,\rho,\gamma) := \rho\left(1+\frac{t^{\gamma}}{\mu}\right)^{-\nu}, \qquad t \geq 0, \tag{19}$$

for $0 < \mu, \nu$ and $0 < \gamma \leq 1$, and where again the use of the parameter ρ will be apparent from the next result. We call $\mathcal{C}_{[0,\pi]}$ the associated restriction to the interval $[0,\pi]$.

Theorem 3 *Let $0 < \gamma \leq 1$. Under the conditions in Lemmas 4 and 5, the matrix-valued function $\mathbf{C}(\cdot) = [C_{ij}(\cdot)]_{i,j=1}^{p}$ defined by*

$$C_{ij}(\theta) = c_{ij}\sigma_{ii}\sigma_{jj}\rho_{ij}\mathcal{C}(\theta;\nu_{ij},\mu_{ij},\rho_{ij},\gamma), \tag{20}$$

belongs to the class Ψ_d^p for any positive integers d and $p \geq 2$.

Proof. Again, the proof is of the constructive type. We have that

$$0 \leq t \mapsto \mathcal{C}(t;\nu,\mu,\rho,\gamma) = \int_{[0,\infty)} \mathrm{e}^{-t^{\gamma}\xi}h(\xi;\nu,\mu,\rho)\mathrm{d}\xi.$$

The proof then comes from noting that the function $t \mapsto \exp(-t^\gamma)$ is completely monotone for any $\gamma \in (0,1]$. Thus, its restriction to the interval $[0,\pi]$ belongs to the class Ψ_d. The scale mixture argument in Theorem A plus the arguments in Lemmas 4 and 5 complete the proof. $\square$

References

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, 180–193.

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

Dai, F., and Xu, Y. (2013), *Approximation Theory and Harmonic Analysis on Spheres and Balls.* Springer monographs in mathematics. Springer, New York.

Daley, D.J. and Porcu, E. (2014), "Dimension walks through Schoenberg spectral measures", *Proceedings of the American Mathematical Society*, 142, 5, 1813–1824

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

Gneiting, T. (2001), "Criteria of Polya Type for Radial Positive Definite Functions," *Proceedings of the American Mathematical Society*, 129, 2309–2318.

Gneiting, T. (2013), "Strictly and Non-Strictly Positive Definite Functions on Spheres," *Bernoulli*, 19, 1327–1349.

Gneiting, T., and Raftery, A. E. (2007), "Strictly Proper Scoring Rules, Prediction, and Estimation," *Journal of the American Statistical Association*, 102, 359–378.

Gneiting, T., and Schlather, M. (2004), "Stochastic Models that Separate Fractal Dimension and the Hurst Effect," *SIAM Review*, 46, 269–282.

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.

Gradshteyn, I. S., and Ryzhik, I. M. (2007), *Table of Integrals, Series, and Products, Seventh edition*, edited by A. Jeffrey and D. Zwillinger, Academic Press.

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

Matérn, B. (1986), *Spatial Variation, 2nd Edition*, Berlin: Springer-Verlag.

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

Porcu, E., and Zastavnyi, V. (2014), "Generalized Askey Functions and Their Walks Through Dimensions," *Expositiones Mathematicae*, 32, 190–198.

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.

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

Schoenberg, I. J. (1942), "Positive Definite Functions on Spheres," *Duke Mathematical Journal*, 9, 96–108.

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.

Zhang, H., and Huang, H.(2010), "Kriging and Cross-Validation for Massive Spatial Data," *Environmetrics*, 21, 290–304.