Stoch Environ Res Risk Assess (2015) 29:193–204
DOI 10.1007/s00477-014-0867-6

ORIGINAL PAPER

Nonstationary matrix covariances: compact support, long range dependence and quasi-arithmetic constructions

William Kleiber · Emilio Porcu

Published online: 23 March 2014

Abstract Flexible models for multivariate processes are increasingly important for datasets in the geophysical, environmental, economics and health sciences. Modern datasets involve numerous variables observed at large numbers of space–time locations, with millions of data points being common. We develop a suite of stochastic models for nonstationary multivariate processes. The constructions break into three basic categories—quasi-arithmetic, locally stationary covariances with compact support, and locally stationary covariances with possible long-range dependence. All derived models are nonstationary, and we illustrate the flexibility of select choices through simulation.

Keywords Compact support · Long range dependence · Matrix-valued covariance · Nonstationary · Quasi-arithmetic functional

1 Introduction

Spatial and spatiotemporal data analysis is a fundamental goal in fields as diverse as statistics, astrophysics, hydrology, ecology, medical geography, environmental and petroleum engineering, remote sensing and geographical information systems (GIS). Modern space–time datasets involve multiple variables observed at between hundreds and millions of locations. The size of these datasets and the intricate nonstationary and cross-process dependence that is commonly present proves to be an insurmountable challenge for the currently available statistical methodology. Herein, we introduce a suite of models for large, complex multivariate spatial datasets that can handle substantial nonstationarity, as well as cross-process dependence.

Throughout this paper we consider m-dimensional (or vector-valued) Gaussian random fields in $d \geq 1$ dimensions, $Z(x) = (Z_1(x), \ldots, Z_m(x))'$, $x \in \mathbb{R}^d$. The assumption of Gaussianity guarantees that, for inference purposes, we only need to consider the second order properties of Z, the mean vector $\mu(x) = \mathbb{E}Z(x)$, the direct covariances $C_{ii}(x, y) = \text{Cov}(Z_i(x), Z_i(y))$ for $i = 1, \ldots, m$, and the cross-covariances $C_{ij}(x, y) = \text{Cov}(Z_i(x), Z_j(y))$ for $i \neq j$. Such mathematical objects form an $m \times m$ matrix of functions $\mathbf{C}(x, y) \in M_{m\times m}$, with (i, j)th component $C_{ij}(x, y)$. The matrix function $\mathbf{C}(x, y)$ represents the covariance structure of $Z(x)$ if and only if C is nonnegative definite in the sense that, for any n-dimensional finite system of m-dimensional vectors $\{a_k\}_{k=1}^n$ and for any n–dimensional collection of locations $\{x_k\}_{k=1}^n$, we have

$$\sum_{i,j=1}^{m} \sum_{k,\ell=1}^{n} a_{ik} C_{ij}(x_k, x_\ell) a_{j\ell} \geq 0.$$

Such a property is difficult to ensure and requires a serious mathematical effort for any candidate matrix-valued function $\mathbf{C}$. The importance of these constructions was recognized many decades ago in a seminal paper by Cramér (1940), who considered stationary constructions $\mathbf{C}(x, y) = \mathbf{C}(\|x - y\|)$, where $\| \cdot \|$ denotes the Euclidean seminorm.

Most of the literature of the last 30 years is based on the assumption of stationarity and isotropy of $\mathbf{C}$, so that $\mathbf{C}(x, y) = \mathbf{C}(\|x - y\|)$. For example, the linear model of coregionalization (LMC; Goulard and Voltz 1992) has

W. Kleiber (✉)
Department of Applied Mathematics, University of Colorado, Boulder, CO 80309, USA
e-mail: william.kleiber@colorado.edu

E. Porcu
Departamento de Matemática, Universidad Técnica Federico Santa Maria, Valparaiso, Chile

been very popular for over twenty years, although its limits are illuminated by Wackernagel (2003) and Gneiting et al. (2010). Very recently, more effort has been devoted to the construction of second order models that describe the dependence over space and time of vector-valued random fields. The easiest way to construct matrix-valued covariances is through separability, where $C_{ij}(\cdot) = a_{ij}C(\cdot)$, for C a valid univariate covariance function and $A = [a_{ij}]_{i,j=1}^m$ a positive definite matrix of coefficients (Mardia and Goodall 1993). The construction is easy to implement but not very interesting in terms of interpretability and flexibility. Kernel and covariance convolution methods have also been very popular in recent years, and they may be useful provided some closed form expression is available through convolution, which is usually the limiting factor (Ver Hoef and Barry 1998; Gaspari and Cohn 1999; Majumdar and Gelfand 2007; Majumdar et al. 2010). Apanasovich and Genton (2010) and Apanasovich et al. (2012) have proposed some interesting constructions through latent processes and through generalizations of the results of Gneiting et al. (2010), respectively. Finally, Porcu and Zastavnyi (2011) provided permissibility criteria that can be used to show that a candidate matrix-valued mapping can be used as correlation function of a vector-valued random field.

When dealing with multivariate random fields, one should distinguish between processes representing vectorial physical variables, e.g., velocity of a moving particle, and multivariate random fields representing a state of vectors, the components of which may have very different magnitudes. The latter class shall be named in this paper *adimensional* multivariate random fields. In the first case, the vector random fields must obey specific mathematical constraints (such as zero divergence or zero curl) that result from physical laws. This is the case of a statistically isotropic second-rank tensor whose spectral density has a specific form (Furutsu 1963). In a similar context, Narcowich and Ward (1994) proposed curl-free and divergence-free matrix-valued kernels, see Scheuerer and Schlather (2012) for recent extensions and discussion. In vector calculus, the curl (or rotor) is a vector operator that describes the infinitesimal rotation of a 3-dimensional vector field, whilst divergence is a vector operator that measures the magnitude of a vector field's source or sink at a given point, in terms of a signed scalar. In particular, let ∇ be the $d \times 1$ gradient vector, $\Delta = \nabla^T \nabla$ be the Laplacian, and $C : \mathbb{R}^d \to \mathbb{R}$ be a positive definite function. Then

$$C_{\text{tra}}(y-x) = (-\Delta I + \nabla\nabla^T)C(y-x) \quad \text{and}$$
$$C_{\text{lon}}(y-x) = (-\nabla\nabla^T)C(y-x)$$

are matrix-valued covariances with zero divergence and zero curl, respectively. Note that in the spectral representation $C_{\text{tra}}(\cdot)$ and $C_{\text{lon}}(\cdot)$ correspond to the transverse and longitudinal term, respectively.

This paper is devoted to nonstationary models for adimensional vector–valued random fields. For these, the crucial scientific question is: what kind of nonstationarity do we want to attain? Recent literature devoted to nonstationary modeling regards what is termed here *locally reducible stationarity*: a nonstationary covariance is obtained either by spatial adaptation (Paciorek and Schervish 2006; Pintore and Holmes 2006; Porcu et al. 2009a; Majumdar et al. 2010; Kleiber and Nychka 2012; Kleiber and Genton 2013), or by convolution techniques through locally stationary kernels (Fuentes and Smith 2001; Fuentes 2002; Higdon 1998; Kleiber et al. 2013). The common denominator amongst these techniques is that stationary covariances are special cases of the nonstationary one (hence, locally reducible stationarity). In particular, Kleiber and Nychka (2012) proposed a locally stationary reducible covariance whose functional form is of the Matérn type. The Matérn is parameterized by a range α and a smoothness ν, where, for $\xi \in [0, \infty)$, we define $\mathcal{M}_{\alpha,\nu}(\xi) = (\alpha\xi)^\nu \mathcal{K}_\nu(\alpha\xi)$ for a modified Bessel function of the second kind of order ν. Extending Paciorek and Schervish (2006) and Stein (2005), the authors propose covariances of the form

$$C_{ij}(x,y) \propto \mathcal{M}_{\alpha_{ij}(x,y),\nu_{ij}(x,y)}(\xi(x,y)) \tag{1}$$

where now $\xi(x, y)$ is the Mahalanobis distance defined by a positive definite matrix function, and $\alpha_{ij}(x,y) = (\alpha_i(x) + \alpha_j(y))/2$ and $\nu_{ij}(x,y) = (\nu_i(x) + \nu_j(y))/2$ are locally adaptive functions. This nonstationary Matérn is locally reducible stationary since the special case $\nu_i(x) = \nu_i$ and $\alpha_i(x) = \alpha_i$ offers the stationary multivariate Matérn model proposed by Gneiting et al. (2010).

It is important to notice that most previous literature on this subject is based on the technique of spatial adaptation of the parameters indexing a parametric family of stationary covariances. Such framework allows, on the one hand, to work with algebraically tractable closed form, but on the other hand implies an ill-posed problem of estimation of such spatially adaptive parameters. Also, normally one must impose restrictions on the parameter functions in order to preserve the validity of the resulting structure, which can lead to further difficulties, depending on the situation at hand.

This paper is inspired by the following goals:

(a) Propose a class of multivariate covariance models whose elements depend separately on the coordinates x and y and additionally is irreducible in the following sense: there are no pairs (R_{ij}, Φ_{ij}), $i, j = 1, \ldots, m$, for bijections on $\mathbb{R}^d$, Φ_{ij}, and $[R]_{ij} =$

$R_{ij}(\cdot)$ a positive definite matrix-valued function, such that

$$C_{ij}(x,y) = R_{ij}\big(\Phi_{ij}(y) - \Phi_{ij}(x)\big), \qquad (x,y) \in \mathbb{R}^d \times \mathbb{R}^d. \tag{2}$$

(b) Propose a class of kernels $\psi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and a collection of m functions $f_i : \mathbb{R}^d \to \mathbb{R}$ such that the composition

$$C_{ij}(x,y) = \psi\big(f_i(x), f_j(y)\big), \qquad (x,y) \in \mathbb{R}^d \times \mathbb{R}^d, \tag{3}$$

offers a positive definite matrix-valued function. Such a class would avoid the problem of using spatially adaptive parameters so that the resulting nonstationary matrix-valued covariance can be estimated through, for instance, maximum likelihood techniques.

(c) Under the framework of locally stationary reducible covariances, there are two lines of discussion that should be mentioned here:

(c.1)
The multivariate locally stationary reducible Matérn model of (1) does not have compact support, and it is desirable to develop a multivariate locally stationary reducible model whose members $C_{ij}(\cdot)$ are compactly supported, say in the unit sphere. This is required particularly for high-dimensional datasets.

(c.2)
On the other hand, the multivariate Matérn has light tails, hence it is desirable to have models indexing long range dependence (also called the Hurst effect).

2 Background and methodology

2.1 Nonnegative definite matrix functions

We start by fixing our notation. Let $M_{m\times m}$ denote the set of all $m \times m$ matrices with entries in $\mathbb{C}$. The matrix $A \in M_{m\times m}$ is nonnegative definite if the inequality $z^{*}Az \geq 0$ holds for every $z \in \mathbb{C}^m$, where z^* is the conjugate transpose.

Let E be a real linear space. A matrix function $\mathbf{C} : E \times E \to M_{m\times m}$ is called nonnegative definite on E if the inequality

$$\sum_{k,\ell=1}^{n} z_k^{*}\mathbf{C}(x_k, x_\ell) z_\ell = \sum_{k,\ell=1}^{n}\sum_{i,j=1}^{m} \bar{z}_{ki} C_{ij}(x_k, x_\ell) z_{\ell j} \geq 0 \tag{4}$$

holds for any finite collection of points $\{x_k\}_{k=1}^n \in E$ and complex vectors $z_1, \ldots, z_n \in \mathbb{C}^m$. The set of all positive definite matrix functions $\mathbf{C} : E \times E \to M_{m\times m}$ is denoted by $\Phi^m(E)$. When $m = 1$ we have $\Phi(E) = \Phi^1(E)$ for the set of all positive definite complex valued functions $f : E \to \mathbb{C}$. In this case, we can give the following equivalent definition: a complex-valued function $f : E \times E \to \mathbb{C}$ is said to belong to the class $\Phi(E)$ if for any finite collection of points $\{x_k\}_{k=1}^n \in E$ the matrix $\left[f(x_k, x_\ell)\right]_{k,\ell=1}^n$ is nonnegative definite,

$$\text{for all } a_1, \ldots, a_n \in \mathbb{C}, \qquad \sum_{k,\ell=1}^{n} \bar{a}_k f(x_k, x_\ell) a_\ell \geq 0$$

we will finally write $\Phi^m = \Phi^m(\mathbb{R}^0)$ for the class of positive-definite matrices.

Throughout the paper we shall make use of completely monotone functions defined on the positive real line, being infinitely often differentiable functions f whose derivatives change sign in the following sense: $(-1)^n f^{(n)}(x) \geq 0$, $\forall\, n \in \mathbb{N}, x > 0$. According to Bernstein's theorem, such functions are the Laplace transform of positive and bounded measures,

$$f(t) = \int_0^\infty e^{-rt}\mu(\mathrm{d}r). \tag{5}$$

2.2 Quasi-arithmetic composition of two real functions

Let Ψ be the class of real functions φ defined in some domain $D(\varphi) \subset \mathbb{R}$, admitting a proper inverse φ^{-1}, defined in $D(\varphi^{-1}) \subset \mathbb{R}$, and such that $\varphi(\varphi^{-1}(t)) = t$ for all $t \in D(\varphi^{-1})$. For now on we shall write $\mathcal{D}$ for any subset of $\mathbb{R}^d$, being normally a compact set.

For $f_1, f_2 : \mathcal{D} \to \mathbb{R}_+$ such that $f_1(\mathcal{D}) \cup f_2(\mathcal{D}) \subset D(\varphi)$, for some $\varphi \in \Psi$, denote by $\mathcal{Q}_\varphi(f_1, f_2)$ the quasi-arithmetic composition of f_1 and f_2 with generating function φ, and define it as

$$\mathcal{Q}_\varphi(f_1, f_2)(x,y) = \varphi^{-1}\left(\frac{1}{2}\varphi \circ f_1(x) + \frac{1}{2}\varphi \circ f_2(y)\right), \quad (x,y) \in \mathcal{D} \times \mathcal{D}, \tag{6}$$

where $\circ$ denotes the composition of two functions. Quasi-arithmetic means have a long history and they can be traced back to Nagumo (1930) and Hardy et al. (1934). Recently, Porcu et al. (2009b) proposed criteria for the permissibility of quasi-arithmetic compositions for scalar-valued random fields and part of this paper is devoted to generalizing their results to the case of vector-valued fields. Note that our definition is indeed equivalent to that in Porcu et al. (2009b), as the function φ is monotonic and admits a proper inverse.

Four basic examples of quasi-arithmetic compositions of functions are shown in Table 1. Some conventions are needed in order to solve possible ill-defined values. We

Table 1 Examples of quasi-arithmetic compositions for some possible choices of the generating function $\varphi \in \Psi$

$\varphi(\xi)$	$\varphi^{-1}(\xi)$	$\mathcal{Q}_{\varphi}(f_1,f_2)\ (x,y)$	Remarks
$\exp(-\xi)$	$-\log \xi$	$f_1(x)^{1/2} f_2(y)^{1/2}$	$f_1,f_2 : \mathcal{D} \to [0,\infty)$ $\log 0 = -\infty$ $\exp(-\infty) = 0$
$1/\xi$	$1/\xi$	$\frac{2f_1(x)f_2(y)}{f_1(x)+f_2(y)}$	$f_1,f_2 : \mathcal{D} \to [0,\infty)$ $1/0 = \infty,\ 1/\infty = 0$ $0/0 = 0$
$M(1-\xi/M)_+$	$M(1-\xi/M)_+$	$\frac{1}{2}f_1(x) + \frac{1}{2}f_2(y)$	$f_1,f_2 : \mathcal{D} \to [0,M]$ for some M > 0 $(u)_+ = \max(u,0)$
$-\log \xi$	$\exp(-\xi)$	$-\log\left(\frac{\exp(-f_1(x)) + \exp(-f_2(y))}{2}\right)$	$f_1,f_2 : \mathcal{D} \to [0,\infty)$

will show in subsequent sections that many other examples of compositions can be obtained but here we remark only those leading to a well known average operator.

3 Multivariate covariances through quasi-arithmetic compositions and Laplace transforms

In this section we propose some techniques for direct construction of (cross)-covariance matrices that satisfy the goals (a) and (b) stated in Sect. 1. The idea is straightforward: use the quasi-arithmetic class of functionals in Eq. (6) as a link function between the margins such that

$$C_{ij}(x,y) = \mathcal{Q}_{\varphi}\big(f_i,f_j\big)(x,y), \qquad (x,y) \in \mathcal{D} \times \mathcal{D}, \tag{7}$$

where $f_i : \mathcal{D} \to \mathbb{R}_+$ is a collection of m nonnegative Lebesgue-measurable functions. As shown subsequently, we do not need any other analytic property for such functions. The first nice feature of such construction is that the margins are of the type $\mathcal{Q}_{\varphi}(f_i,f_i)(x,x) = f_i(x)$. The following result offers sufficient conditions for the permissibility of this new candidate class of matrix-valued covariance mappings. A slight additional flexibility added here is the presence of the co-located correlation coefficients ρ_{ij}, which represent the same-site correlations between the ith and jth processes.

Theorem 1 *Let $f_i : \mathcal{D} \to \mathbb{R}_+$ be a set of nonnegative-valued mappings. Let $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ be such that its proper inverse is completely monotone on the positive real line and such that $f_i(\mathcal{D}) \cup f_j(\mathcal{D}) \subset D(\varphi)$ for all pairs $i,j = 1,\ldots,m$. Define the mapping $\mathbf{C} : \mathcal{D} \times \mathcal{D} \to M_{m\times m}$ as*

$$\mathbf{C}(x,y) = \big[C_{ij}(x,y)\big]_{i,j=1}^{m} = \big[\sigma_i\sigma_j\rho_{ij}\mathcal{Q}_{\varphi}\big(f_i,f_j\big)(x,y)\big]_{i,j=1}^{m}, \qquad (x,y) \in \mathcal{D} \times \mathcal{D}, \tag{8}$$

where ρ_{ij} is a co-located correlation coefficient such that the matrix $[\rho_{ij}]_{i,j=1}^{m}$ is nonnegative definite, and $\sigma_i \geq 0$. Then $\mathbf{C}$ is a valid multivariate covariance.

Proof We give a proof of the constructive type. We need to prove that, for any finite collection of points $\{x_i\}_{i=1}^{N}$ of $\mathcal{D}$ and vectors $\{c_i\}_{i=1}^{N}$ of $\mathbb{C}^m$, the following inequality holds,

$$\sum_{k,\ell=1}^{N}\sum_{i,j=1}^{m} c_{ik}\overline{c}_{j\ell}\sigma_i\sigma_j\mathcal{Q}_{\varphi}\big(f_i,f_j\big)(x_k,x_\ell) \geq 0$$

where we do not need to include ρ_{ij} since its matrix is nonnegative definite by assumption, which is preserved under Schur products (Bhatia 2007). In particular,

$$\begin{aligned}
&\sum_{k,\ell=1}^{N}\sum_{i,j=1}^{m} c_{ik}\overline{c}_{j\ell}\sigma_i\sigma_j\mathcal{Q}_{\varphi}\big(f_i,f_j\big)(x_k,x_\ell)\\
&= \sum_{k,\ell=1}^{N}\sum_{i,j=1}^{m} c_{ik}\overline{c}_{j\ell}\sigma_i\sigma_j\varphi^{-1}\left(\frac{1}{2}\varphi\circ f_i(x_k) + \frac{1}{2}\varphi\circ f_j(x_\ell)\right)\\
&= \sum_{k,\ell=1}^{N}\sum_{i,j=1}^{m} c_{ik}\overline{c}_{j\ell}\sigma_i\sigma_j\int_{[0,\infty)}\exp\left(-\frac{r}{2}\varphi\circ f_i(x_k) - \frac{r}{2}\varphi\circ f_j(x_\ell)\right)\mu(\mathrm{d}r)\\
&= \int_{[0,\infty)}\left|\sum_{k=1}^{N}\sum_{i=1}^{m}\sigma_i c_{ik}\exp\left(-\frac{r}{2}\varphi\circ f_i(x_k)\right)\right|^2\mu(\mathrm{d}r) \geq 0
\end{aligned}$$

where the third equality comes from Bernstein's theorem, where μ is a positive bounded measure on $\mathbb{R}_+$. □

A huge quantity of examples can be proposed under this setting. We do not need the completely monotone function to be finite at the origin (as required, for instance, in Gneiting 2002b). Consider functions f_i being radial in their argument, in the sense that there exist functions $\psi_i : \mathbb{R} \to \mathbb{R}_+$ such that

$$f_i(x) = \psi_i(\|x\|), \qquad x \in \mathbb{R}^d,$$

where $\|\cdot\|$ can be any seminorm. Another possibility is to compose the functions ψ_i with the great circle distance; this would not affect the construction proposed in Theorem 1.

Table 2 Examples of completely monotone functions

Function φ^{-1}	Parameter restrictions	Function φ^{-1}	Parameter restrictions
$(1+x^{\alpha})^{\beta}$	$0<\alpha\leq 1, \beta<0$	$x^{\nu}\mathcal{K}_{\nu}(x)$	$\nu>0$
$\left(\frac{x^{\beta}}{1+x^{\beta}}\right)^{\gamma}$	$0<\beta\leq 1, 0<\gamma<1$	$\left(\exp(\sqrt{t})+\exp(-t\sqrt{t})\right)^{-\nu}$	$\nu>0$

Table 3 Examples of complete Bernstein functions

Function	Parameter restrictions	Function	Parameter restrictions
$1-\frac{1}{(1+x^{\alpha})^{\beta}}$	$0<\alpha, \beta\leq 1$	$e^{x}-x\left(1+\frac{1}{x}\right)^{x}-\frac{x}{x+1}$	
$\left(\frac{x^{\rho}}{1+x^{\rho}}\right)^{\gamma}$	$0<\gamma, \rho<1$	$\frac{1}{a}-\frac{1}{x}\log\left(1+\frac{x}{a}\right)$	$a>0$
$\frac{x^{\alpha}-x(1+x)^{\alpha-1}}{(1+x)^{\alpha}-x^{\alpha}}$	$0<\alpha<1$	$\sqrt{\frac{x}{2}}\frac{\sinh^{2}\sqrt{2x}}{\sinh(2\sqrt{2x})}$	
$\sqrt{x}\left(1-e^{-2a\sqrt{x}}\right)$	$a>0$	$x^{1-\nu}e^{ax}\,\Gamma(\nu;ax)$	$a>0,\ 0<\nu<1$
$\frac{x\left(1-e^{-2\sqrt{x+a}}\right)}{\sqrt{x+a}}$	$a>0$	$x^{\nu}e^{a/x}\,\Gamma(\nu;\frac{a}{x})$	$a>0,\ 0<\nu<1$

$\Gamma(a;x)=\int_x^{\infty}t^{a-1}e^{-t}\,dt$ is the incomplete Gamma function

Regarding the choice of the generator φ^{-1} for the composition $\mathcal{Q}_{\varphi}$ in Eq. (6), some completely monotone functions are listed in Gneiting (2002b) and we report them together with other choices in Table 2.

The fact that the family of completely monotone functions is not very rich can be overcome by considering completely Bernstein functions; the excellent textbook by Schilling et al. (2010) offers a wide selection. We point the reader to their book as well as to Porcu and Schilling (2011) for an historical account of the use of such functions (under different names) over several branches of mathematics. Table 3 is taken directly from Porcu and Schilling (2011). Such functions are very important since they have stability properties that make them appealing in order to create new examples of completely monotone functions. The Stieltjies class of functions (Berg and Forst 1975) is a subclass of the completely monotonic class. Let us denote with $\mathcal{CBF}$ and $\mathcal{S}$, respectively, the class of completely Bernstein and Stieltjies functions. Using the arguments in Porcu and Schilling (2011), we have, for any $f\not\equiv 0$,

$$f\in\mathcal{CBF}\iff\left[x\to\frac{f(x)}{x}\right]\in\mathcal{S}\iff\left[x\to\frac{x}{f(x)}\right]\in\mathcal{CBF}\iff\frac{1}{f}\in\mathcal{S}, \tag{9}$$

so that, using simple stability properties, we have a substantial class of completely monotone functions that can be used as generator of the composition $\mathcal{Q}_{\varphi}$ in Eq. (6). The same functions can then be effectively used as functions ψ_i, $i=1,\ldots,m$, entering the composition in Eq. (7).

Consider covariances that can be obtained through the choice $\varphi^{-1}(t)=t^{-\delta}$, $\delta>0$. For instance, if $f_i(x)=(1+\|x\|^{\alpha_i})_+^{-\beta_i}$, $\alpha_i,\beta_i>0$, we get a nonstationary covariance of the type

$$C_{ij}(x,y)=\left(1/2(1+\|x\|^{\alpha_i})_+^{\beta_i/\delta}+1/2(1+\|y\|^{\alpha_j})_+^{\beta_j/\delta}\right)^{-\delta}, \tag{10}$$

having the interesting property that the marginal variances $C_{ii}(x,x)=f_i(x)$, which are radial functions of the generalized Cauchy type (Gneiting and Schlather 2004) and thus the variances of the processes Z_i at a point $x\in\mathbb{R}^d$ are decreasing and convex on the positive real line. This may be desirable for some processes, but if the opposite were desired, then it would be sufficient to apply the decomposition (10) to the function $g_i(x)=1-f_i(x)$ to obtain increasing variances. Finally, constant variance can be obtained by $C_{ij}(x,y)/\sqrt{C_{ii}(x,x)C_{jj}(y,y)}$, yielding still a permissible matrix-valued covariance function. The covariance (10) is not readily reduced to a stationary covariance; it is a nonstationary construction with straightforward parameterization. We anticipate a primary application for (10) to be modeling dispersion from a source, such as particulate matter from a volcano, or wind vectors from a hurricane's eye.

Many other examples can be obtained under the same setting. For instance, taking $f_i(x)=\|x\|^{\beta\gamma}(1+\|x\|^{\beta})^{\gamma}$, $\beta\in(0,2]$ and $\gamma\in(0,1]$ we respect the requirements in Theorem 1 since such function is bounded by one and thus we obtain, for $\delta=\gamma_i$, $i=1,\ldots,m$,

$$C_{ij}(x,y)=\left(\frac{1}{2}\|x\|^{-\beta_i}(1+\|x\|^{\beta_i})+\frac{1}{2}\|y\|^{-\beta_j}(1+\|y\|^{\beta_j})\right)^{-\delta}, \tag{11}$$

and $C_{ij}(0,0)=0$, which has increasing marginal variances. The proposed structure is nonstationary and irreducible in the sense of point (b) of the introduction. The following result is obtained using much similar arguments as in Porcu et al. (2010).

Theorem 2 *Let φ be a strictly monotonic function and $\mathbf{C}(x,y)$ be the matrix of functions as defined in Eq. (8). If there exists an even mapping $R_{ij}:\mathbb{R}^d\to\mathbb{R}$ and a bijection Φ_{ij} such that the reducibility condition (2) holds, then f_i, f_j and R_{ij} are constant functions.*

The Gneiting class of space–time correlation functions (Gneiting 2002b; Porcu and Zastavnyi 2011) has been widely used in applications involving space–time data. For φ a completely monotone function and h an increasing and concave function, such a correlation is defined as

$$C(x-y,t-s)=\frac{1}{h^{d/2}\left(|t-s|^2\right)}\varphi\left(\frac{\|x-y\|^2}{h\left(|t-s|^2\right)}\right). \qquad (12)$$

Gneiting (2002b) describes sufficient conditions for such function to be the stationary covariance associated with a space–time Gaussian random field. Porcu and Zastavnyi (2011) examine necessary conditions, and relax the hypothesis on the function h, which is restricted to a function whose first derivative is completely monotonic in Gneiting (2002b). Porcu and Zastavnyi (2011) also analyze how to preserve permissibility if the function φ is not composed with the Euclidean norm, but with an arbitrary seminorm.

It is well known that completely monotone functions are the Laplace transforms of nonnegative and bounded measures. The natural generalization is thus to consider bivariate Laplace transforms $\mathcal{L}(\cdot,\cdot)$ associated with a random vector, admitting the integral representation, for $(\xi_1,\xi_2)\in\mathbb{R}_+^2$,

$$\mathcal{L}(\xi_1,\xi_2)=\int\limits_{[0,\infty)}\int\limits_{[0,\infty)}\exp(-r_1\xi_1-r_2\xi_2)\mu(\mathrm{d}r_1,\mathrm{d}r_2) \qquad (13)$$

where μ is a nonnegative measure on $\mathbb{R}_+^2$.

This allows us to generalize the Gneiting class to the nonstationary case. We omit the proof since it will be obtained following the same arguments as in Theorem 1.

Theorem 3 *Let $\mathcal{L}$ be the Laplace transform of a positive bivariate random vector. Let $\psi_{ki}:\mathbb{R}_+\to\mathbb{R}_+$ $k=1,2$ and $i=1,\ldots,m$ be m-dimensional collection of Lebesgue measurable functions. Then*

$$C_{ij}(x,t,y,u)=\frac{\sigma_i\sigma_j\rho_{ij}}{\psi_{1i}(t^2)^{d/2}\psi_{1j}(u^2)^{d/2}}\times\mathcal{L}\left(\psi_{2i}\left(\frac{\|x\|^2}{\psi_{1i}(t^2)}\right),\psi_{2j}\left(\frac{\|y\|^2}{\psi_{1j}(u^2)}\right)\right), \qquad (14)$$

for $(x,t,y,u)\in\mathcal{D}\times\mathbb{R}\times\mathcal{D}\times\mathbb{R}$, is a permissible matrix-valued space–time covariance function.

The Laplace transform for the Frechet-Hoeffding lower bound of bivariate copula has expression

$$\mathcal{L}(\xi_1,\xi_2)=\frac{\exp^{-\xi_1}-\exp^{-\xi_2}}{\xi_2-\xi_1}.$$

We may now choose the following functions: $\psi_{1i}(t)=1+t^{\alpha_i}$, $\alpha_i>0$, $\psi_{2i}(t)=t^{\delta_i}$ in a way to obtain

$$C_{ij}(x,t,y,u)=\frac{\sigma_i\sigma_j}{(1+t^{\alpha_i})^{d/2}(1+u^{\alpha_j})^{d/2}}\frac{\exp\left(\frac{\|x\|^2}{1+t^{\alpha_i}}\right)^{\delta_i}-\exp\left(\frac{\|y\|^2}{1+u^{\alpha_j}}\right)^{\delta_j}}{\frac{\|y\|^2}{1+u^{\alpha_j}}-\frac{\|x\|^2}{1+t^{\alpha_i}}},$$

and $C_{ij}(0,t,0,u)=\sigma_i\sigma_j/((1+t^{\alpha_i})^{d/2}(1+u^{\alpha_j})^{d/2})$, having the nice feature of being asymmetric in both time instants u and t.

Before moving to the next section, we cover the proof of Theorem 3.

Proof Suppose there are m processes; consider a finite collection of space–time coordinates, $(x_i,t_i), i=1,\ldots,N$, and arbitrary vectors $\{c_i\}_{i=1}^N$ of $\mathbb{C}^m$. As in the proof of Theorem 1, we do not need to include ρ_{ij} as its matrix is assumed to be nonnegative definite, which is preserved under Schur products (Bhatia 2007).

Then,

$$\begin{aligned}
&\sum_{k,\ell=1}^N\sum_{i,j=1}^m c_{ik}\overline{c}_{j\ell}\sigma_i\sigma_jC_{ij}(x_k,t_k,x_\ell,t_\ell)\\
&=\sum_{k,\ell=1}^N\sum_{i,j=1}^m c_{ik}\overline{c}_{j\ell}\frac{\sigma_i\sigma_j}{\psi_{1i}(t_k^2)^{d/2}\psi_{1j}(t_\ell^2)^{d/2}}\\
&\quad\times\mathcal{L}\left(\psi_{2i}\left(\frac{\|x_k\|^2}{\psi_{1i}(t_k^2)}\right),\psi_{2j}\left(\frac{\|x_\ell\|^2}{\psi_{1j}(t_\ell^2)}\right)\right)\\
&=\sum_{k,\ell=1}^N\sum_{i,j=1}^m c_{ik}\overline{c}_{j\ell}\frac{\sigma_i\sigma_j}{\psi_{1i}(t_k^2)^{d/2}\psi_{1j}(t_\ell^2)^{d/2}}\\
&\quad\times\int\limits_{[0,\infty)}\int\limits_{[0,\infty)}\exp\left(-r_1\psi_{2i}\left(\frac{\|x_k\|^2}{\psi_{1i}(t_k^2)}\right)\right.\\
&\quad\left.-r_2\psi_{2j}\left(\frac{\|x_\ell\|^2}{\psi_{1j}(t_\ell^2)}\right)\right)\mu(\mathrm{d}r_1,\mathrm{d}r_2)\\
&=\sum_{k,\ell=1}^N\sum_{i,j=1}^m\int\limits_{[0,\infty)}\int\limits_{[0,\infty)}\frac{c_{ik}\sigma_i}{\psi_{1i}(t_k^2)^{d/2}}\exp\left(-r\psi_{2i}\left(\frac{\|x_k\|^2}{\psi_{1i}(t_k^2)}\right)\right)\\
&\quad\times\frac{\overline{c}_{j\ell}\sigma_j}{\psi_{1j}(t_\ell^2)^{d/2}}\exp\left(-r\psi_{2j}\left(\frac{\|x_\ell\|^2}{\psi_{1j}(t_\ell^2)}\right)\right)\mu(\mathrm{d}r_1,\mathrm{d}r_2)\\
&=\int\limits_{[0,\infty)}\int\limits_{[0,\infty)}\left|\sum_{k=1}^N\sum_{i=1}^m\frac{c_{ik}\sigma_i}{\psi_{1i}(t_k^2)^{d/2}}\exp\left(-r\psi_{2i}\left(\frac{\|x_k\|^2}{\psi_{1i}(t_k^2)}\right)\right)\right|^2\\
&\quad\mu(\mathrm{d}r)\geq 0
\end{aligned}$$

where the third equality comes from Bernstein's theorem, where μ is a positive bounded measure on $\mathbb{R}_+$.

4 Locally stationary covariances with compact support

Modern spatial datasets typically involve multiple processes at thousands to tens of thousands of spatial locations. Traditional geostatistical constructions are not well adapted to such scenarios; indeed covariance matrices with large dimensions are either infeasible or impossible to invert, and thereby precludes traditional likelihood and kriging ventures. While a number of solutions have been proposed, particularly for kriging, using compactly supported covariances has proven an effective idea, either for direct use, or for tapering a non-compact covariance (Furrer et al. 2006; Kaufman et al. 2008; Du et al. 2009). While some authors have recently acknowledged the need for such constructions (Du and Ma 2012; Porcu et al. 2013a), there is yet a lack of flexible nonstationary possibilities. In this section we examine compactly supported matrix covariances, and in particular derive a class of such models that allow for substantial nonstationarity.

4.1 Wendland–Gneiting functions

For an exposition of our following constructions, we start by describing a popular class of functions in the statistical and numerical analysis literature, proposed by Wendland (1995) in the numerical analysis setting and then by Gneiting (2002a) in the geostatistical one. This Wendland-Gneiting class of correlation functions has been repeatedly used in applications involving, for example, the so-called tapered likelihood (Furrer et al. 2006). Let

$$\psi_{\nu,0}(t) = (1-t)_+^{\nu}, \qquad t \geq 0, \quad \nu \in \mathbb{R}_+, \tag{15}$$

be the truncated power function, also known as the Askey function (Askey 1973). We make similar use of ψ_ν or $\psi_{\nu,0}$ as will be apparent from the context; we have that $x \mapsto \psi_\nu(\|x\|)$, $x \in \mathbb{R}^d$, is compactly supported over the unit sphere in $\mathbb{R}^d$, and $\psi_\nu \in \Phi_d$ for $\nu \geq (d+1)/2$. For any $g \in \Phi(\mathbb{R}^d)$ for which $\lim_{t\to\infty} \int_0^t u\, g(u)\, \mathrm{d}u < \infty$, the Descente operator I of Matheron (1962) is defined by

$$Ig(t) = \frac{\int_t^\infty u\, g(u)\, \mathrm{d}u}{\int_0^\infty u\, g(u)\, \mathrm{d}u} \qquad (t \in \mathbb{R}_+).$$

Wendland (1995) defines

$$\psi_{d,k}(t) = I^k \psi_{[\frac{1}{2}d]+k+1,0}(t), \qquad t \geq 0, \tag{16}$$

via k-fold iterated application of the Descente operator on the Askey function $\psi_{\nu,0}(x)$ defined at (15), and proves that $\psi_{d,k} \in \Phi(\mathbb{R}^d)$. The implications in terms of differentiability are well summarized by Gneiting (2002a): $\psi_{d,k}$ is a polynomial of order $[\frac{1}{2}d] + 3k + 1$ and differentiable of order $2k$ on $\mathbb{R}$. Moreover, $\psi_{d,k} \in C^{2k}(\mathbb{R})$ are unique up to a constant factor, and the polynomial degree is minimal for given space dimension d and smoothness $2k$; that is, the degree of the piecewise polynomials is minimal for the given smoothness and dimension for which the radial basis function should be positive definite.

4.2 Results: compact support

We start with the Askey function $\psi_\nu(\cdot)$ of (15). The following theorem characterizes a large class of compactly supported covariances, and after its proof we discuss the problem of differentiability at the origin.

Theorem 4 *Suppose $\gamma_{ij}(x,y) = (\gamma_i(x) + \gamma_j(y))/2$ are positive valued mappings for $i = 1, \ldots, m$. Define the matrix-valued mapping $\mathbf{C} : \mathbb{R}^d \times \mathbb{R}^d \to M_{m\times m}$ with*

$$C_{ij}(x,y) = \begin{cases} b^{\nu+1} B(\gamma_{ij}(x,y)+1, \nu+1)\, \psi_{\nu+\gamma_{ij}(x,y)+1}\left(\dfrac{\|x-y\|}{b}\right), & x,y \in \mathbb{R}^d, \\ 0, & \textit{otherwise}. \end{cases} \tag{17}$$

where B is the beta function. Then, $\mathbf{C}$ is a nonnegative definite matrix-valued mapping.

Proof Using notation and $\gamma_{ij} = (\gamma_i + \gamma_j)/2$ for the mappings defined in the assertion, we have that the function $f(t;x,y) = t^\nu (1 - t/b)_+^{\gamma_{ij}}$ is nonnegative definite on $\mathbb{R}^{2d}$ for any fixed positive t, as is the function $\psi_\nu(\|x-y\|/t)$ for the previously defined arguments. From Theorem 1 in Porcu and Zastavnyi (2011), we thus have that the scale mixture integral

$$C_{ij}(x,y) = \int_{\mathbb{R}_+} \psi_\nu\left(\frac{\|x-y\|}{t}\right) t^\nu \left(1 - \frac{t}{b}\right)_+^{\gamma_{ij}} \mathrm{d}t$$

offers a nonnegative definite matrix function. In particular, we have trivially that $C_{ij}(x,y) < \infty$, and that $\psi_\nu(\|\cdot\|/t)$ is

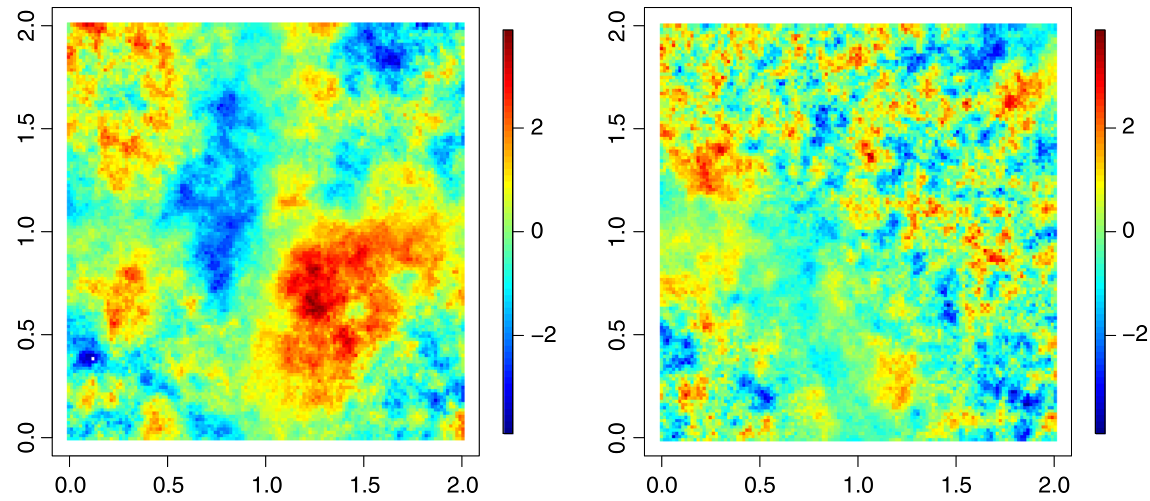

Fig. 1 Compactly supported bivariate simulation. The first variable (*left panel*) has shorter length scale near the four corners, while the second variable has longer length scale in a swath crossing the west to south domain boundaries. The two processes are positively cross-correlated and nonstationary

nonnegative definite (Gneiting 2002a). That the matrix-valued function whose (i,j)th entry is $t^{\nu}\left(1-\frac{t}{b}\right)_{+}^{\gamma_{ij}}$ is a nonnegative definite matrix function follows since it can be written $t^{\nu}\mathbf{f}\mathbf{f}'$ where the ith entry of $\mathbf{f}$ is $\left(1-\frac{t}{b}\right)_{+}^{\gamma_i/2}$, and outer products are nonnegative definite (Bhatia 2007); thus, the conditions (i)–(iii) of Theorem 1 of Porcu and Zastavnyi (2011) hold. Direct inspection then shows that $C_{ij}(x,y)$ can be written as

$$\int_{\|x-y\|}^{b}\left(1-\frac{\|x-y\|}{t}\right)^{\nu}t^{\nu}\left(1-\frac{t}{b}\right)_{+}^{\gamma_{ij}}\mathrm{d}t$$

$$=b^{-\gamma_{ij}}\int_{\|x-y\|}^{b}(t-\|x-y\|)^{\nu}(b-t)_{+}^{\gamma_{ij}}\mathrm{d}t$$

$$=b^{-\gamma_{ij}}\int_{0}^{b-\|x-y\|}(z)^{\nu}(b-\|x-y\|-z)_{+}^{\gamma_{ij}}\mathrm{d}t$$

which gives (17) through integration by parts.

The stationary case of this theorem has been proposed in (Porcu et al. 2013a). Figure 1 illustrates the flexibility of this nonstationary construction, where we have two positively correlated spatial processes, each with distinct and drastically changing marginal nonstationarity. The Descente operator can then be used to obtain new constructions based on (17). For instance, direct calculations (Porcu et al. 2013b) show that

$$B(\gamma+1,\nu+2k+1)I^{k}\psi_{\nu+\gamma+1}(t)$$
$$=\int t^{\nu+2k}(1-t/b)_{+}^{\gamma}I^{k}\psi_{\nu}\left(\frac{\|x-y\|}{t}\right)\mathrm{d}t,$$

and thus the mapping

$$C_{ij}(x,y)=\begin{cases} b^{\nu+2k+1}B(\gamma_{ij}(x,y)+1,\nu+2k+1)\psi_{\nu+\gamma_{ij}(x,y)+1,k}\left(\frac{\|x-y\|}{b}\right), & x,y\in\mathbb{R}^{d},\\ 0, & \text{otherwise}, \end{cases}$$

is a valid model under the same conditions as in Theorem 4 for $\nu=[\frac{1}{2}n]+k+2$. For instance, for $k=1$ we obtain

$$C_{ij}(x,y)=\begin{cases} b^{\nu+3}B(\gamma_{ij}(x,y)+1,\nu+3)\left(1-\frac{\|x-y\|}{b}\right)^{\nu+\gamma_{ij}(x,y)+1}\left(1+(\gamma_{ij}(x,y)+\nu+1)\frac{\|x-y\|}{b}\right)\\ 0, \quad \text{otherwise}. \end{cases}$$

5 Locally stationary covariances with long range dependence

An alternative approach to building matrix covariances is to use normal scale mixtures (Schlather 2010). The following theorem combines quasi arithmetic compositions with normal scale mixtures to produce a general class of nonstationary matrix-valued covariance functions.

For the following theorem, set $\Sigma_{ij}(x,y) = (\Sigma_i(x) + \Sigma_j(y))/2$, where Σ_i maps to the set of real valued positive definite $d \times d$ dimensional matrices, and let $\sigma_i : \mathbb{R}^d \to [0,\infty)$ for $i,j = 1,\ldots,m$. For a measure space $(\Omega, \mathcal{A}, \mu)$ and Borel measurable functions $\theta_i : \mathbb{R}^d \to \mathbb{R}$, $i = 1,2$, the following theorem considers functions $f_i : \mathcal{A} \times \mathbb{R}$ such that $f_i(\cdot, \theta_i(x))$ is measurable with respect to μ for a given $x \in \mathbb{R}^d$. A slight change of notation is also needed here. For the quasi-arithmetic composition of functions f_i, we use $\mathcal{Q}(f_i(\omega;\theta_i(\cdot)), f_j(\omega;\theta_j(\cdot))(x,y)$.

Theorem 5 *Let $f_i : \mathcal{A} \times \mathbb{R}$ as above, with $i = 1,\ldots,m$. For given $x,y \in \mathbb{R}^d$, suppose $\mathcal{Q}_\varphi(f_i(\omega;\theta_i(\cdot)), f_j(\omega;\theta_j(\cdot)))(x,y) \in L^2(\mu)$ for some nonnegative measure μ on $[0,\infty)$ and $i,j = 1,\ldots,m$, for some generator φ. Then the matrix-valued function with (i,j)th entry $C_{ij}(x,y)$ defined as*

$$\frac{\sigma_i(x)\sigma_j(y)}{|\Sigma_{ij}(x,y)|^{1/2}} \int_0^\infty \exp\left(-\omega(x-y)'\Sigma_{ij}(x,y)^{-1}(x-y)\right) \mathcal{Q}_\varphi(f_i(\omega;\theta_i(\cdot)), f_j(\omega;\theta_j(\cdot)))(x,y)\, \mathrm{d}\mu(\omega)$$

is a multivariate covariance function.

Proof Suppose we have m processes indexed by $i,j = 1,\ldots,m$, n locations $x_k, x_\ell \in \mathbb{R}^d, k,\ell = 1,\ldots,n$, and an arbitrary vector $a = (a_{11}, a_{12}, \ldots, a_{mn})'$. Then let Σ be an $mn \times mn$ block matrix made up of m^2, $n \times n$ blocks. Set the (i,j)th block to be an $n \times n$ matrix whose (k,ℓ)th entry is $C_{ij}(x_k, x_\ell)$. The following argument shows $a'\Sigma a \geq 0$. We drop the local standard deviation functions $\sigma_i(x)$ from the proof, as these trivially do not affect the nonnegative definiteness of the resulting matrix Σ. First note the covariance functions can be written

$$\begin{aligned} C_{ij}(x_k,x_\ell) &= |\Sigma_{ij}(x_k,x_\ell)|^{-1/2} \int \exp\left(-\omega(x_k-x_\ell)'\Sigma_{ij}(x_k,y_\ell)^{-1}(x_k-x_\ell)\right) \\ &\quad \times \mathcal{Q}_\varphi(f_i(\omega;\theta(\cdot)), f_j(\omega;\theta(\cdot)))(x_k,x_\ell)\, \mathrm{d}\mu(\omega) \\ &= |\Sigma_{ij}(x_k,x_\ell)|^{-1/2} \int \int_0^\infty \exp\left(-\omega(x_k-x_\ell)'\Sigma_{ij}(x_k,x_\ell)^{-1}(x_k-x_\ell)\right) \\ &\quad \times g_i(\omega;x_k,r) g_j(\omega;x_\ell,r) \mathrm{d}\mu(\omega)\mathrm{d}\xi(\mathrm{r}) \\ &= \int \int_0^\infty 4^{-1/2}(4\pi)^{-d/2}\omega^{-1/2} \int_{\mathbb{R}^d} K_{ik}^\omega(u) K_{j\ell}^\omega(u) \mathrm{d}\eta(\mathrm{u}) \mathrm{g_i}(\omega;\mathrm{x_k},\mathrm{r}) \\ &\quad \times g_j(\omega;x_\ell,r)\mathrm{d}\mu(\omega)\mathrm{d}\xi(\mathrm{r}) \end{aligned}$$

where in the second equality we have made used of Bernstein representation for completely monotonic functions and have used the notation $g_i(\omega;x,r)$ for $\exp(-r\varphi \circ f_i(\omega;\theta(x)))$. Here, $K_{ik}^\omega(u)$ is a Gaussian kernel with mean x_k and variance $\Sigma_i(x_k)/(4\omega)$; see Paciorek and Schervish (2006) for the univariate case. With the above representation, we can write

$$\begin{aligned} a'\Sigma a &= \sum_{i,j=1}^m \sum_{k,\ell=1}^n a_{ik}a_{j\ell} C_{ij}(x_k,x_\ell) \\ &= \sum_{i,j=1}^m \sum_{k,\ell=1}^n a_{ik}a_{j\ell} 2^{-1}(4\pi)^{-d/2} \int \int_0^\infty \int_{\mathbb{R}^d} \omega^{-1/2} K_{ik}^\omega(u) K_{j\ell}^\omega(u) \mathrm{d}\eta(u) \\ &\quad \times g_i(\omega;x_k,r) g_j(\omega;x_\ell,r)\mathrm{d}\mu(\omega)\mathrm{d}\xi(\mathrm{r}) \\ &= 2^{-1}(4\pi)^{-d/2} \int \int_0^\infty \int_{\mathbb{R}^d} \left(\sum_{i=1}^m \sum_{k=1}^n \omega^{-1/4} a_{ik} K_{ik}^\omega(u) g_i(\omega;x_k,r) \right)^2 \\ &\quad \mathrm{d}\eta(u)\mathrm{d}\mu(\omega)\mathrm{d}\xi(r) \geq 0. \end{aligned}$$

□

Theorem 5 is a general construction for nonstationary multivariate covariance functions. The nonstationary multivariate Matérn construction of Kleiber and Nychka (2012) is a special case of Theorem 5, where $\mathcal{Q}_\varphi(f_1(\omega;\theta_1(\cdot)), f_2(\omega;\theta_2(\cdot))(x,y) = f_1(\omega)^{1/2} f_2(\omega)^{1/2}$, $\mathrm{d}\mu(\omega) = \omega^{-1}\exp(-1/(4\omega))$ and $f_i(\omega;\nu_i(x)) = \omega^{-\nu_i(x)}$. In this construction, Σ_i is the locally varying geometric anisotropy, σ_i

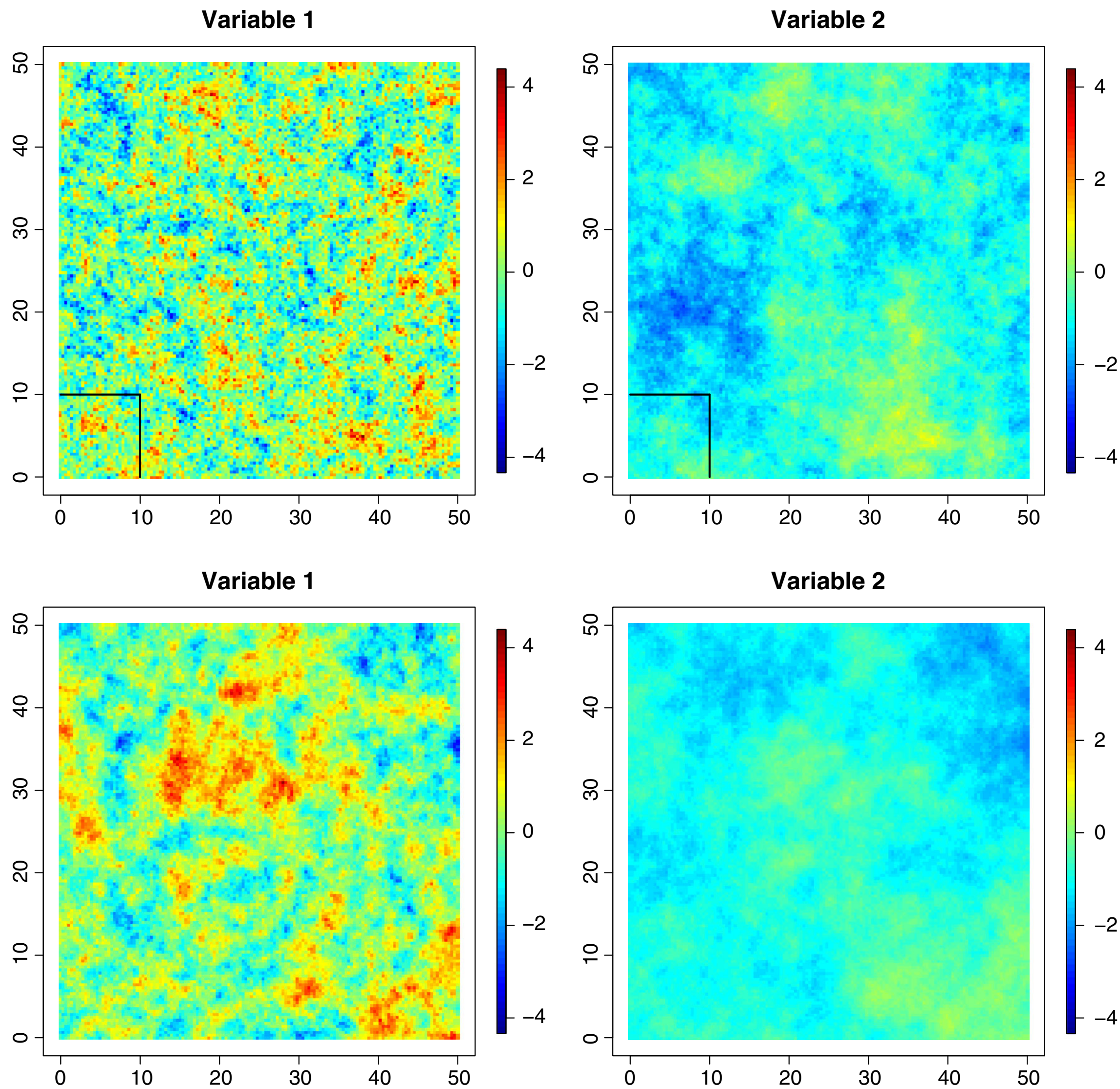

Fig. 2 Realization of a stationary bivariate Gaussian process with the first variable (*left column*) exhibiting short range dependence, while the second variable (*right column*) exhibits long range dependence. The *bottom row* is a detailed zoom-in of the delineated subdomain of the *top row*. The two processes are positively cross-correlated and stationary

is the local standard deviation and v_i is the local smoothness function for the ith process.

Consider an alternative special case of Theorem 5, where $\mathcal{A}_\psi(f_1,f_2)(\omega) = f_1(\omega)^{1/2} f_2(\omega)^{1/2}$. Set $f_i(\omega;x) = \omega^{\delta_i(x)}$ and set the integration measure to $\mathrm{d}\mu(\omega) = \omega^{-1}e^{-\omega}\mathrm{d}\omega$. These choices yield a multivariate covariance function with (i,j)th entry

$$C_{ij}(x,y) = \frac{\sigma_i(x)\sigma_j(y)}{|\Sigma_{ij}(x,y)|^{1/2}(1+(x-y)'\Sigma_{ij}(x,y)^{-1}(x-y))^{(\delta_i(x)+\delta_j(y))/2}}. \tag{18}$$

In the stationary case, $\delta_i(x) = \delta_i$ controls the range-dependence of the process. In particular, with $\Sigma_i(x) = \Sigma_i$ and $\sigma_i(x) = \sigma_i$, then the corresponding process is long-range dependent if $\delta_i \leq d/2$, and is short-range dependent otherwise. This class allows for nontrivial cross-covariance between processes that are long-range dependent, short-range dependent, or some mix thereof. In particular, to the authors' knowledge, this is the first time a class of matrix-valued covariance functions has been described where one process is long-range dependent, another is short-range dependent and there is nontrivial cross-covariance between the two.

Figures 2 and 3 illustrate the flexible mixture of long and short-range dependence endowed by (18). Two positively correlated variables are displayed; in Fig. 2 the two are stationary, with the first being short-range dependent, while the second is long-range dependent. Figure 3

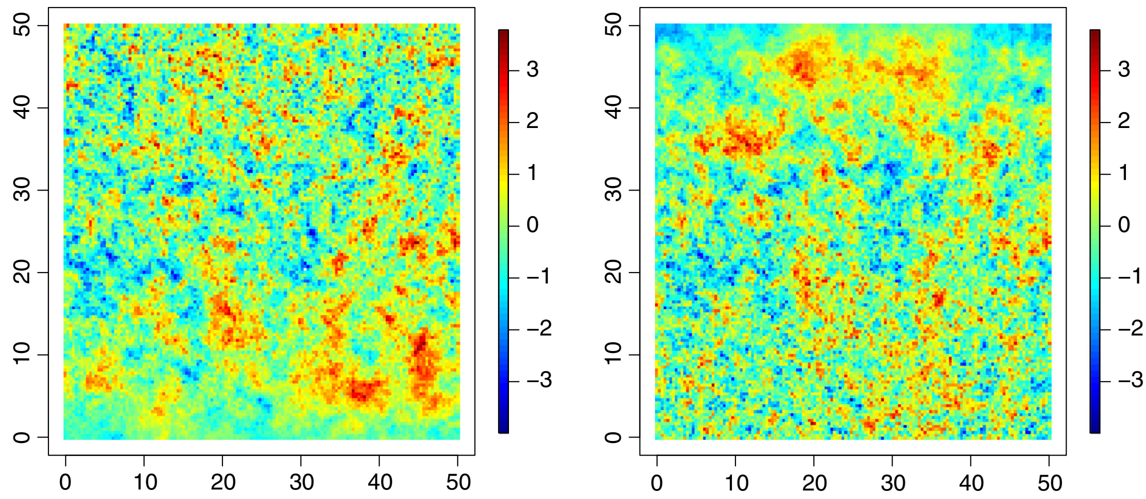

Fig. 3 Realization of a nonstationary bivariate Gaussian process with the first variable exhibiting long-range spatial dependence on southern domain edge, smoothly decaying to short-range dependence on the north domain edge. The second process exhibits long and short range dependence in the opposite direction. The two processes are positively cross-correlated and nonstationary

contains a bivariate simulation with both variables being nonstationary, and exhibiting a mixture of short and long-range dependence across the simulation domain, while still having substantial positive cross-correlation.

Acknowledgments Emilio Porcu is supported by Proyecto Fondecyt Regular number 1130647, funded by the Chilean Ministry of Education.

References

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

Apanasovich TV, Genton MG, Sun Y (2012) A valid Matérn class of cross-covariance functions for multivariate random fields with any number of components. J Am Stat Assoc 107:180–193

Askey R (1973) Radial characteristic functions, Tech. Rep. 1262. Mathematical Research Center, University of Wisconsin, Madison

Berg C, Forst G (1975) Potential theory on locally compact abelian groups. Springer, Berlin

Bhatia R (2007) Positive definite matrices. Princeton Press, Princeton, New Jersey, USA

Cramér H (1940) On the theory of stationary random processes. Ann Math 41:215–230

Du J, Ma C (2012) Vector random fields with compactly supported covariance matrix functions. J Stat Plan Inference 143:457–467

Du J, Zhang H, Mandrekar VS (2009) Fixed-domain asymptotic properties of tapered maximum likelihood estimators. Ann Stat 37:3330–3361

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

Fuentes M, Smith RL (2001) A new class of nonstationary spatial models, Tech. rep. North Carolina State University, Department of Statistics, Raleigh, NC

Furrer R, Genton MG, Nychka D (2006) Covariance tapering for interpolation of large datasets. J Comput Gr Stat 15:502–523

Furutsu K (1963) On the theory of radio wave propagation over inhomogeneous earth. J Res Natl Bureau Stand 67D:39–62

Gaspari G, Cohn SE (1999) Construction of correlation functions in two and three dimensions. Q J R Meteorol Soc 125: 723–757

Gneiting T (2002a) Compactly supported correlation functions. J Multivar Anal 83:493–508

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

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

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

Goulard M, Voltz M (1992) Linear coregionalization model: tools for estimation and choice of cross-variogram matrix. Math Geol 24:269–282

Hardy GH, Littlewood JE, Pólya G (1934) Inequalities. Cambridge University Press, Cambridge

Higdon D (1998) A process-convolution approach to modelling temperatures in the North Atlantic Ocean. Environ Ecol Stat 5:173–190

Kaufman CG, Schervish MJ, Nychka DW (2008) Covariance tapering for likelihood-based estimation in large spatial data sets. J Am Stat Assoc 103:1545–1555

Kleiber W, Genton MG (2013) Spatially varying cross-correlation coefficients in the presence of nugget effects. Biometrika 100: 213–220

Kleiber W, Nychka D (2012) Nonstationary modeling for multivariate spatial processes. J Multivar Anal 112:76–91

Kleiber W, Katz RW, Rajagopalan B (2013) Daily minimum and maximum temperature simulation over complex terrain. Ann Appl Stat 7:588–612

Majumdar A, Gelfand AE (2007) Multivariate spatial modeling for geostatistical data using convolved covariance functions. Math Geol 39:225–245

Majumdar A, Paul D, Bautista D (2010) A generalized convolution model for multivariate nonstationary spatial processes. Stat Sin 20:675–695

Mardia K, Goodall C (1993) Spatial-temporal analysis of multivariate environmental monitoring data. In: Patil GP, Rao CR (eds)

Multivariate environmental statistics. North Holland, Amsterdam, pp 347–386
Matheron G (1962) Traité de géostatistique appliquée. Tome 1, Editions Technip, Paris
Nagumo M (1930) Über eine klasse der mittelwerte. Jpn J Math 7:71–79
Narcowich FJ, Ward JD (1994) Generalized Hermite interpolation via matrix-valued conditionally positive definite functions. Math Comput 63:661–687
Paciorek CJ, Schervish MJ (2006) Spatial modelling using a new class of nonstationary covariance functions. Environmetrics 17:483–506
Pintore A, Holmes C (2006) Spatially adaptive non-stationary covariance functions via spatially adaptive spectra
Porcu E, Schilling RL (2011) From Schoenberg to Pick–Nevanlinna: toward a complete picture of the variogram class. Bernoulli 17:441–455
Porcu E, Zastavnyi V (2011) Characterization theorems for some classes of covariance functions associated to vector valued random fields. J Multivar Anal 102:1293–1301
Porcu E, Gregori P, Mateu J (2009a) Archimedean spectral densities for nonstationary space–time geostatistics. Stat Sin 19:273–286
Porcu E, Mateu J, Christakos G (2009b) Quasi-arithmetic means of covariance functions with potential applications to space–time data. J Multivar Anal 100:1830–1844
Porcu E, Matkowski J, Mateu J (2010) On the non-reducibility of non-stationary correlation functions to stationary ones under a class of mean-operator transformations. Stoch Environ Res Risk Assess 24:599–610
Porcu E, Daley DJ, Buhmann M, and Bevilacqua M (2013a) Radial basis functions with compact support for multivariate geostatistics. Stoch Environ Res Risk Assess 27(4):909–922
Porcu E, Daley DJ, and Bevilacqua M (2013b) Classes of compactly supported correlation functions for multivariate random fields
Scheuerer M, Schlather M (2012) Covariance models for divergence-free and curl-free random vector fields. Stoch Models 28:433–451
Schilling RL, Song R, Vondraček Z (2010) Bernstein functions: theory and. Springer, Berlin
Schlather M (2010) Some covariance models based on normal scale mixtures. Bernoulli 16:780–797
Stein ML (2005) Nonstationary spatial covariance functions. University of Chicago, CISES Technical Report 21
Ver Hoef JM, Barry RP (1998) Constructing and fitting models for cokriging and multivariable spatial prediction. J Stat Plan Inference 69:275–294
Wackernagel H (2003) Multivariate geostatistics, 3rd edn. Springer, Berlin
Wendland H (1995) Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. Adv Comput Math 4:389–396