Bernoulli **16**(3), 2010, 780–797
DOI: 10.3150/09-BEJ226

Some covariance models based on normal scale mixtures

MARTIN SCHLATHER

Institut für Mathematische Stochastik & Zentrum für Statistik, Universität Göttingen, Goldschmidtstr. 7, D-37077 Göttingen, Germany. E-mail: schlather@math.uni-goettingen.de

Modelling spatio-temporal processes has become an important issue in current research. Since Gaussian processes are essentially determined by their second order structure, broad classes of covariance functions are of interest. Here, a new class is described that merges and generalizes various models presented in the literature, in particular models in Gneiting (*J. Amer. Statist. Assoc.* **97** (2002) 590–600) and Stein (Nonstationary spatial covariance functions (2005) Univ. Chicago). Furthermore, new models and a multivariate extension are introduced.

Keywords: cross covariance function; Gneiting's class; rainfall model; spatio-temporal model

1. Introduction

Spatio-temporal modelling is an important task in many disciplines of the natural sciences, geosciences, and engineering. Hence, the development of models for spatio-temporal correlation structure is of particular interest. The lively activity in this field of research has become apparent through various recent reviews of known classes of spatio-temporal covariance functions (Gneiting *et al.* (2007), Mateu *et al.* (2008), Ma (2008)). To categorise these classes, different aspects have been considered. Gneiting *et al.* (2007) distinguish between the properties of covariance functions, such as motion invariance, separability, full symmetry, or conformity with Taylor's hypothesis. Another classification is based on the construction principles (Ma (2008)), such as spectral methods (Stein (2005a)), multiplicative mixture models (Ma (2002)), additive models (Ma (2005c)), turning bands upgrade (Kolovos *et al.* (2004)), derivatives and integrals (Ma (2005b)), and Gneiting's (2002) approach, see also Stein (2005c) and Ma (2003).

Surprisingly, some rather different approaches to the construction of spatial and spatio-temporal covariance models can be subsumed in a unique class of normal scale mixtures, which is a generalization of Gneiting's (2002) class. As its construction is based on cross covariance functions, Section 2 illustrates some of the properties of cross covariance functions and cross variograms. In Section 3, Gneiting's class itself is generalized. Section 4 introduces two new classes of spatio-temporal models. Section 5 presents an extension to multivariate models. In addition to the two-dimensional realisations illustrated below, three-dimensional realisations are available in the form of films at the following website: www.stochastik.math.uni-goettingen.de/data/bernoulli10/.

2. Background: Cross covariance functions

Here we introduce some basic notions and properties of cross covariance functions and cross variograms. See Wackernagel (2003) for a geostatistical overview and Reisert and Burkhardt (2007) for some of the construction principles of multivariate cross covariance functions in a general framework.

Let $Z(x) = (Z_1(x), \dots, Z_m(x))$, $x \in \mathbb{R}^d$, be a zero mean, second order m-variate, complex valued random field in $\mathbb{R}^d$, that is, $\operatorname{Var} Z_j(x)$ exists and $\mathbb{E}Z_j(x) = 0$ for all $x \in \mathbb{R}^d$ and $j = 1, \dots, m$. Then, the cross covariance function $C : \mathbb{R}^{2d} \to \mathbb{C}^{m\times m}$ is defined by

$$C_{jk}(x, y) = \operatorname{Cov}(Z_j(x), Z_k(y)), \qquad x, y \in \mathbb{R}^d, \ j, k = 1, \dots, m.$$

Clearly $C(x, y) = \overline{C^\top(y, x)}$, but $C(x, y) = \overline{C^\top(x, y)}$ is not valid in general. A function $C : \mathbb{R}^{2d} \to \mathbb{C}^{m\times m}$ with $C(x, y) = \overline{C^\top(y, x)}$, $x, y \in \mathbb{R}^d$, is called positive definite if for all $n \in \mathbb{N}$, $x_1, \dots, x_n \in \mathbb{R}^d$ and $a_1, \dots, a_n \in \mathbb{C}^m$,

$$\sum_{p=1}^{n} \sum_{q=1}^{n} a_p^\top C(x_p, x_q) \bar{a}_q \geq 0. \tag{1}$$

It is called strictly positive definite if strict inequality holds in (1) for $(a_1, \dots, a_n) \neq 0$ and pairwise distinct $x_1, \dots, x_n$. Accordingly, we name a Hermitian matrix $M \in \mathbb{C}^{m\times m}$ positive definite, if $v^\top M \bar{v} \geq 0$ for all $v \in \mathbb{C}^m$, and strictly positive definite if strict inequality holds for $v \neq 0$.

As in the univariate case, we derive from Kolmogorov's existence theorem that a function $C : \mathbb{R}^{2d} \to \mathbb{C}^{m\times m}$ with $C(x, y) = \overline{C^\top(y, x)}$ is a positive definite function if and only if a (Gaussian) random field exists with C as cross covariance function. Further, a function $C : \mathbb{R}^{2d} \to \mathbb{R}^{m\times m}$ is a positive definite function if and only if Equation (1) holds for any $a_1, \dots, a_n \in \mathbb{R}^m$.

The cross variogram $\gamma : \mathbb{R}^{2d} \to \mathbb{C}^{m\times m}$, $\gamma = (\gamma_{jk})_{j,k=1,\dots,m}$ is defined by

$$\gamma_{jk}(x, y) = \tfrac{1}{2}\mathbb{E}\big(Z_j(x) - Z_j(y)\big)\overline{\big(Z_k(x) - Z_k(y)\big)}, \qquad x, y \in \mathbb{R}^d, \ j, k = 1, \dots, m.$$

If Z has second order stationary increments, then $\gamma(x, y)$ depends only on the distance vector $h = x - y$, that is, $\gamma(x, y) = \tilde{\gamma}(h)$ for some function $\tilde{\gamma} : \mathbb{R}^d \to \mathbb{C}^{m\times m}$. If in addition Z is univariate, then $\tilde{\gamma}$ is called a (semi-)variogram. Schoenberg's (1938b) theorem states that a function $\tilde{\gamma} : \mathbb{R}^d \to \mathbb{R}$ with $\tilde{\gamma}(0) = 0$ is a variogram if and only if $\exp(-r\tilde{\gamma})$ is a covariance function for all $r > 0$, see also Gneiting *et al.* (2001). Let us now discuss multivariate and non-stationary versions of this statement. To this end, we denote the componentwise multiplication of matrices by "$*$", in particular,

$$A^{*n} = (A_{jk}^n)_{jk} \qquad \text{for } A = (A_{jk})_{jk}.$$

Further, $f^*(A)$ denotes the componentwise function evaluation, for example,

$$\exp^*(A) = (\exp(A_{jk}))_{jk}.$$

Theorem 1. *Let $C : \mathbb{R}^{2d} \to \mathbb{C}^{m\times m}$ and $E_{m\times m}$ be the $m \times m$ matrix whose components are all* 1.

1. *The following three assertions are equivalent*: (i) C *is a cross covariance function*; (ii) $\exp^*(rC) - E_{m\times m}$ *is a cross covariance function for all* $r > 0$; (iii) $\sinh^*(rC)$ *is a cross covariance function for all* $r > 0$.
2. *If* $\exp^*(rC)$ *is a cross covariance function for all* $r > 0$ *then*

$$C^{(z)}(x, y) = C(z, z) - C(x, z) - C(z, y) + C(x, y) \tag{2}$$

is a cross covariance function for all $z \in \mathbb{R}^d$. *If* $m = 1$ *and* (2) *holds for one* $z \in \mathbb{R}^d$, *then* $\exp(rC)$ *is a covariance function for all* $r > 0$.

Proof. Note that the componentwise product $C_1 * C_2$ of two m-variate cross covariance functions C_1 and C_2 is again a cross covariance function. To see this, consider the componentwise product of two independent random fields with cross covariance functions C_1 and C_2. In particular, $C(x, y)^{*n}$ and $rC(x, y)$, $r \geq 0$, are cross covariance functions. Furthermore, the sum and the pointwise limit of m-variate cross covariance functions are cross covariance functions.

1. Both functions, $\exp(x) - 1$ and $\sinh(x)$, have Taylor expansion on $\mathbb{R}$ with positive coefficients only. Hence, $\exp^*(rC) - E_{m\times m}$ and $\sinh^*(rC)$ are cross covariance functions if C is a cross covariance function. On the other hand, since the Taylor expansions equal $x + o(x)$ as $x \to 0$, we have that $(\exp^*(rC) - E_{m\times m})/r$ and $\sinh^*(rC)/r$ converge to C as $r \to 0$ and C must be a cross covariance function.

2. The proof follows the lines in Matheron (1972). Let $a_1, \dots, a_n \in \mathbb{C}^m$, $x_1, \dots, x_n \in \mathbb{R}^d$, $a_0 = -\sum_{p=1}^n a_p$ and $x_0 = z$ for some $z \in \mathbb{R}^d$. Then

$$0 \leq \lim_{r\to 0} \sum_{p=0}^n \sum_{q=0}^n a_p^\top \frac{\exp^*(rC(x_p, x_q)) - E_{m\times m}}{r} \bar{a}_q = \sum_{p=0}^n \sum_{q=0}^n a_p^\top C(x_p, x_q)\bar{a}_q$$

$$= \sum_{p=1}^n \sum_{q=1}^n a_p^\top [C(x_p, x_q) + C(z, z) - C(x_p, z) - C(z, x_q)]\bar{a}_q.$$

Conversely, assume that $m = 1$ and Equation (2) holds. Since $C_0(x, y) = f(x)\overline{f(y)}$ is a covariance function for any function $f : \mathbb{R}^d \to \mathbb{C}$ (Berlinet and Thomas-Agnan (2004), Lemma 1) part 1 of the theorem results in

$$\exp(rC(x, y)) = f(x)\overline{f(y)} \exp\bigl(rC(x, y) + rC(z, z) - rC(x, z) - rC(z, y)\bigr)$$

being a positive definite function for any $r > 0$ and $f(x) = \exp(rC(x, z) - rC(z, z)/2)$. □

Remark 2. If $m = 1$, $C(x, y) = -\tilde{\gamma}(x - y)$ and $z = 0$, then $C^{(0)}$ in Equation (2) equals the covariance function of an intrinsically stationary random field Z with $Z(0) = 0$ almost surely, that is, part 2 of Theorem 1 yields Schoenberg's (1938b) theorem. If $m > 1$, the reverse statement in part 2 of Theorem 1 does not hold in general, as the following example shows. Let $M \in \mathbb{R}^{m\times m}$, $m \geq 2$, be a symmetric, strictly positive definite matrix with identical diagonal elements,

$\tilde{\gamma}:\mathbb{R}^d \to \mathbb{R}$ a variogram, and $C(x, y) = -M\tilde{\gamma}(x - y)$. Then $C^{(0)}(x, y)$ given by (2) is a cross covariance function, but $\exp^*(-M\tilde{\gamma})$ is a positive definite function if and only if $\tilde{\gamma} \equiv 0$. To see this, assume that $\exp^*(-M\tilde{\gamma})$ is a positive definite function and let $m = 2$, $M = (M_{jk})_{j,k=1,2}$, and $Z(x) = (Z_1(x), Z_2(x))$ be a corresponding random field. Then with $a = (1, -1, 1, -1)^\top$ we have

$$\begin{aligned}\operatorname{Var}\big(Z_1(0) - Z_2(0) + Z_1(y) - Z_2(y)\big) &= a^\top \begin{pmatrix} \exp^*(-M\tilde{\gamma}(0)) & \exp^*(-M\tilde{\gamma}(y)) \\ \exp^*(-M\tilde{\gamma}(y)) & \exp^*(-M\tilde{\gamma}(0)) \end{pmatrix} a \\ &= 2(1, -1)\exp^*(-M\tilde{\gamma}(y))(1, -1)^\top \\ &= 4\big(\mathrm{e}^{-M_{11}\tilde{\gamma}(y)} - \mathrm{e}^{-M_{12}\tilde{\gamma}(y)}\big).\end{aligned}$$

Since $M_{11} > M_{12}$, the latter is non-negative if and only if $\tilde{\gamma}(y) = 0$.

So, for an arbitrary cross variograms $\gamma:\mathbb{R}^{2d} \to \mathbb{C}^{m\times m}$ the function $\exp^*(-\gamma(x, y))$ is not a positive definite, in general. However,

$$C_1(x, y) = \exp^*\big(\gamma(x, 0) + \gamma(y, 0) - \gamma(x, y)\big)$$

and

$$\begin{aligned} C_2(x, y) &= \exp^*\big(\gamma(x, 0) + \gamma(y, 0) - D_{xy} - \gamma(x, y)\big), \\ (D_{xy})_{jk} &= \gamma_{jj}(x, 0) + \gamma_{kk}(y, 0), \end{aligned} \tag{3}$$

are always positive definite functions in $\mathbb{R}^d$, cf. Theorem 2.2 in Berg *et al.* (1984) for the univariate case. To see this, let γ be an m-variate cross variogram and Z a corresponding m-variate random field. Let $Y(x) = Z(x) - Z(0)$ and $c(x, y) = \mathbb{E}Y(x)Y^\top(y)$. Then c and $\overline{c^\top}$ are positive definite functions and

$$\begin{aligned} c_{jk}(x, y) + \overline{c_{kj}(x, y)} &= \mathbb{E}\big(Y_j(x)\overline{Y_k(y)} + \overline{Y_k(x)}Y_j(y)\big) \\ &= \mathbb{E}\big[Y_j(x)\overline{Y_k(x)} + Y_j(y)\overline{Y_k(y)} + \big(Y_j(x) - Y_j(y)\big)\big(\overline{Y_k(y)} - \overline{Y_k(x)}\big)\big] \\ &= \gamma_{jk}(x, 0) + \gamma_{jk}(y, 0) - \gamma_{jk}(x, y). \end{aligned}$$

Part 1 of Theorem 1 yields that C_1 is a positive definite function. Let Z be a corresponding random field. Then the random field $(\mathrm{e}^{-\gamma_{11}(x,0)}Z_1(x), \ldots, \mathrm{e}^{-\gamma_{mm}(x,0)}Z_m(x))$, $x \in \mathbb{R}^d$, has cross covariance function C_2.

Remark 3. Let $C(x_1, x_2) = VD(x_1, x_2)\bar{V}^\top \in \mathbb{C}^{m\times m}$, $x, y \in \mathbb{R}^d$, for some unitary matrix $V \in \mathbb{C}^{m\times m}$. The values of the mapping $D:\mathbb{R}^{2d} \to \mathbb{C}^{m\times m}$ are diagonal matrices,

$$D(x_1, x_2) = \operatorname{diag}(D_1(x_1, x_2), \ldots, D_m(x_1, x_2)), \qquad x_1, x_2 \in \mathbb{R}^d,$$

and the $D_j:\mathbb{R}^{2d} \to \mathbb{C}$, $j = 1, \ldots, m$, are arbitrary functions. Then the n-fold matrix product $C^n:\mathbb{R}^{2d} \to \mathbb{C}^{m\times m}$ is a cross covariance function in $\mathbb{R}^d$ for any $n \in \mathbb{N}$ if and only if the D_j are all covariance functions, and Theorem 1 remains true if $\exp^*(rC(x,y))$ is replaced by

$$\exp(rC(x,y)) = \sum_{n=0}^{\infty} \frac{r^n C^n(x,y)}{n!}, \qquad x, y \in \mathbb{R}^d.$$

The subsequent proposition generalizes the results in Cressie and Huang (1999) and Theorem 1 in Gneiting (2002). Denote by $\mathcal{B}^d$ the ensemble of Borel sets of $\mathbb{R}^d$.

Proposition 4. *Let d and l be non-negative integers with $d + l > 0$ and $C : \mathbb{R}^{l+2d} \to \mathbb{C}^{m\times m}$ a continuous function in the first argument. Then the following two assertions are equivalent*:

1. *C is a cross covariance function that is translation invariant in the first argument, that is, $C(h, y_1, y_2) = \mathrm{Cov}(Z(x+h, y_1), Z(x, y_2))$ for some second order random field Z on $\mathbb{R}^{l+d}$ and all $x, h \in \mathbb{R}^l$ and $y_1, y_2 \in \mathbb{R}^d$.*
2. *$C : \mathbb{R}^l \times \mathbb{R}^{2d} \to \mathbb{C}^{m\times m}$ is the Fourier transform of some finite measures $F_{y_1,y_2,j,k}$, $y_1, y_2 \in \mathbb{R}^d$, $j, k = 1, \dots, m$, that is,*

$$C_{jk}(h, y_1, y_2) = \int \mathrm{e}^{-\mathrm{i}\langle h,\omega\rangle} F_{y_1,y_2,j,k}(\mathrm{d}\omega), \qquad h \in \mathbb{R}^l,\ j, k = 1, \dots, m, \tag{4}$$

and

$$(C^A_{jk}(y_1, y_2))_{jk} = (F_{y_1,y_2,j,k}(A))_{jk}, \qquad y_1, y_2 \in \mathbb{R}^d, \tag{5}$$

is an m-variate cross covariance function in $\mathbb{R}^d$ for any $A \in \mathcal{B}^l$.

Proof. The proof follows the lines in Gneiting (2002). Let us first assume that Equations (4) and (5) hold. Let $n \in \mathbb{N}$, $x_1, \dots, x_n \in \mathbb{R}^l$, $y_1, \dots, y_n \in \mathbb{R}^d$ and $a_1, \dots, a_n \in \mathbb{C}^m$ be fixed. Then a matrix-valued function $f : \mathbb{R}^{l+2d} \to \mathbb{C}^{m\times m}$ and a non-negative finite measure F on $\mathbb{R}^d$ exists, such that

$$\int_A f_{jk}(\omega, y_p, y_q) F(\mathrm{d}\omega) = F_{y_p,y_q,j,k}(A), \qquad p, q = 1, \dots, n,\ j, k = 1, \dots, m, \tag{6}$$

for any $A \in \mathcal{B}^l$. For instance, let $F(A) = \sum_{p=1}^n \sum_{k=1}^m F_{y_p,y_p,k,k}(A)$. Then, Equation (5) implies that the $mn \times mn$ matrix $(f_{jk}(\omega, y_p, y_q))_{j,k;p,q}$ is hermitian for F-almost all ω. Now,

$$\sum_{p=1}^n \sum_{q=1}^n a_p^\top C(x_p - x_q, y_p, y_q)\overline{a_q} = \int \sum_{p=1}^n \sum_{q=1}^n \mathrm{e}^{-\mathrm{i}\langle x_p,\omega\rangle} a_p^\top f(\omega, y_p, y_q)\overline{\mathrm{e}^{-\mathrm{i}\langle x_q,\omega\rangle}}\,\overline{a_q} F(\mathrm{d}\omega) \geq 0.$$

Conversely, let $C(h, y_1, y_2) : \mathbb{R}^{l+2d} \to \mathbb{C}^{m\times m}$ be a covariance function that is stationary in its first argument. We have

$$C_{jk}(h, y, y') = \int \mathrm{e}^{-\mathrm{i}\langle \omega,h\rangle} F_{y,y',j,k}(\mathrm{d}\omega), \qquad h \in \mathbb{R}^l;\ y, y' \in \mathbb{R}^d,\ j, k = 1, \dots, m,$$

for some finite, not necessarily positive measures $F_{y,y',j,k}$ (Yaglom (1987b), page 115). It now remains to demonstrate that equality (5) holds. Fix $n \in \mathbb{N}$, $y_1, \ldots, y_n \in \mathbb{R}^d$, and $a_1, \ldots, a_n \in \mathbb{C}^m$. Then a non-negative finite measure F and a function $f : \mathbb{R}^{l+2d} \to \mathbb{C}^{m\times m}$ exist, such that Equation (6) holds. By assumption, $\sum_{p=1}^n \sum_{q=1}^n a_p^\top C(\cdot, y_p, y_q)\overline{a_q}$ is a positive definite, continuous function and its Fourier transform is non-negative. Following directly from the linearity of the Fourier transform, we have that for F-almost all $\omega \in \mathbb{R}^l$

$$\sum_{p=1}^n \sum_{q=1}^n a_p^\top f(\omega, y_p, y_q)\overline{a_q} \ge 0,$$

which finally leads to Equation (5). □

If a covariance function is translation invariant, we will write only one argument for ease of notation, for example, $C(h)$, $h = x - y \in \mathbb{R}^d$, instead of $C(x, y)$, $x, y \in \mathbb{R}^d$.

3. Generalized Gneiting's class

A function $C(x, y) = \varphi(\|h\|)$, $h = x - y \in \mathbb{R}^d$, is a motion invariant, real-valued covariance function in $\mathbb{R}^d$ for all $d \in \mathbb{N}$ if and only if φ is a normal scale mixture, that is,

$$\varphi(h) = \int_{[0,\infty)} \exp(-ah^2)\,\mathrm{d}F(a), \qquad h \ge 0,$$

for some non-negative measure F (Schoenberg (1938a)). Examples are the stable model (Yaglom (1987a)), the generalized Cauchy model (Gneiting and Schlather (2004)),

$$\varphi(h) = (1 + h^\alpha)^{-\beta/\alpha}, \qquad h \ge 0,$$

$\alpha \in [0, 2]$, $\beta > 0$, and the generalized hyperbolic model (Barndorff-Nielsen (1979), Gneiting (1997)). The latter includes as special case the Whittle–Matérn model (Stein (1999)),

$$\varphi(h) = W_\nu(h) = 2^{1-\nu}\Gamma(\nu)^{-1} h^\nu K_\nu(h), \qquad h > 0.$$

Here, $\nu > 0$ and K_ν is a modified Bessel function.

Theorem 5. *Assume that m and d are positive integers and $H : \mathbb{R}^d \to \mathbb{R}^m$. Suppose that φ is a normal scale mixture and $G : \mathbb{R}^{2d} \to \mathbb{R}^{m\times m}$ is a cross variogram in $\mathbb{R}^d$ or $-G$ is a cross covariance function. Let $M \in \mathbb{R}^{m\times m}$ be positive definite, such that $M + G(x, y)$ is strictly positive definite for all $x, y \in \mathbb{R}^d$. Then*

$$C(x, y) = \frac{\varphi([(H(x) - H(y))^\top (M + G(x, y))^{-1}(H(x) - H(y))]^{1/2})}{\sqrt{|M + G(x, y)|}}, \qquad x, y \in \mathbb{R}^d, \tag{7}$$

is a covariance function in $\mathbb{R}^d$.

Lemma 6. *Let $\gamma : \mathbb{R}^{2d} \to \mathbb{C}^{m\times m}$ be a cross variogram (cross covariance function) in $\mathbb{R}^d$ and $A \in \mathbb{C}^{l\times m}$. Then $\gamma_0 = A\gamma \overline{A^\top}$ is an l-variate, cross variogram (cross covariance function) in $\mathbb{R}^d$.*

Proof of Theorem 5. We follow the proof in Gneiting (2002) but assume first that $\varphi(h) = \mathrm{e}^{-h^2}$. If $G(x, y)$ is a cross variogram, then, according to Lemma 6,

$$g(x, y) = \omega^\top G(x, y)\omega$$

is a (univariate) variogram for any $\omega \in \mathbb{R}^m$. Equation (3) or Theorem 2.2 in Berg *et al.* (1984) implies

$$C_\omega(x, y) = \exp(-\omega^\top G(x, y)\omega), \qquad x, y \in \mathbb{R}^d, \tag{8}$$

and hence,

$$\hat{C}(\omega, x, y) = \exp\bigl(-\omega^\top\bigl(M + G(x, y)\bigr)\omega\bigr), \qquad x, y \in \mathbb{R}^d, \tag{9}$$

are both covariance functions for any fixed $\omega \in \mathbb{R}^m$. With $\mathrm{d}F_{x,y,1,1}(\omega) = \hat{C}(\omega, x, y)\,\mathrm{d}\omega$, Proposition 4 yields that the univariate function

$$C(h, x, y) = c\frac{\exp(-h^\top(M + G(x, y))^{-1}h)}{\sqrt{|M + G(x, y)|}}, \qquad h \in \mathbb{R}^m; x, y \in \mathbb{R}^d$$

is a covariance function in $\mathbb{R}^{m+d}$ for all $c \geq 0$, which is translation invariant in the first argument. Now, consider a random field $Z(\zeta, x)$ on $\mathbb{R}^{m+d}$ corresponding to $C(h, x, y)$ with $c = 1$. Define the random field Y on $\mathbb{R}^d$ by

$$Y(x) = Z(H(x), x).$$

Then the covariance function of Y is equal to the covariance function given in the theorem. For general φ, the assertion is obtained directly from the definition of normal scale mixtures. In case $-G$ is a cross covariance function, the proof runs exactly the same way. □

Example 7. A well known construction of a cross covariance function in $\mathbb{R}^d$ used in machine learning is

$$\tilde{G}(x, y) = f(x)f(y)^\top, \qquad x, y \in \mathbb{R}^d,$$

for some function $f : \mathbb{R}^d \to \mathbb{R}^{m\times l}$. Assume that $M - f(x)f(y)^\top$ is strictly positive definite for all x and y and some positive definite matrix M. Then, C in Equation (7) is a covariance function with $G = -\tilde{G}$.

We denote by $\mathbf{1}_{d\times d} \in \mathbb{R}^{d\times d}$ the identity matrix.

Example 8. Gneiting (2002) delivers a rather general construction of non-separable models based on completely monotone functions, containing as particular case the models developed by Cressie and Huang (1999). Let φ be a completely monotone function, that is, $\varphi(t^2)$, $t \in \mathbb{R}$, is a normal scale mixture, and ψ be a positive function with a completely monotone derivative. Then

$$C(h,u) = \frac{1}{\psi(|u|^2)^{d/2}} \varphi\big(\|h\|^2/\psi(|u|^2)\big), \qquad h \in \mathbb{R}^d, u \in \mathbb{R}, \tag{10}$$

is a translation invariant covariance function in $\mathbb{R}^{d+1}$ (Gneiting (2002), Theorem 2). According to Bernstein's theorem, the function $\psi(\|\cdot\|^2) - c$ is a variogram for some positive constant c, see also Berg *et al.* (1984). The positive definite nature of C in (10) is also ensured by Theorem 5 for $m = d$ and $G((x_1, x_2), (y_1, y_2)) = \psi(\|x_2 - y_2\|^2)\mathbf{1}_{d\times d}$, $x_1, y_1 \in \mathbb{R}^d$, $x_2, y_2 \in \mathbb{R}$. Gneiting (2002) provides examples for ψ and, along the way, introduces a new class of variograms,

$$\gamma(h) = (\|h\|^a + 1)^b - 1, \qquad a \in (0,2], b \in (0,1].$$

This class generalizes the class of variograms of fractal Brownian motion and that of multiquadric kernels (Wendland (2005)).

Example 9. In the context of modelling rainfall, Cox and Isham (1988) proposed in $\mathbb{R}^{d+1}$ the translation invariant covariance function

$$C(h,u) = \mathbb{E}_V \varphi(\|h - Vu\|), \qquad h \in \mathbb{R}^d, u \in \mathbb{R}.$$

Here, $\varphi(\|\cdot\|)$ is a motion invariant covariance function in $\mathbb{R}^d$ and V is a d-dimensional random wind speed vector. Unfortunately, this appealing model has lacked explicit representations. Now let us assume that V follows a d-variate normal distribution $\mathcal{N}(\mu, D/2)$ and $\varphi(x) = \exp(-x^2)$. Then,

$$C(h,u) = \frac{1}{\sqrt{|\mathbf{1}_{d\times d} + u^2 D|}} \varphi\big([(h-u\mu)^\top (\mathbf{1}_{d\times d} + u^2 D)^{-1}(h-u\mu)]^{1/2}\big), \qquad h \in \mathbb{R}^d, u \in \mathbb{R},$$

please refer to the appendix for a proof. Hence, $C(h,u)$ above is a covariance function for any normal mixture φ. Figure 1 provides realizations of a random field with the above covariance

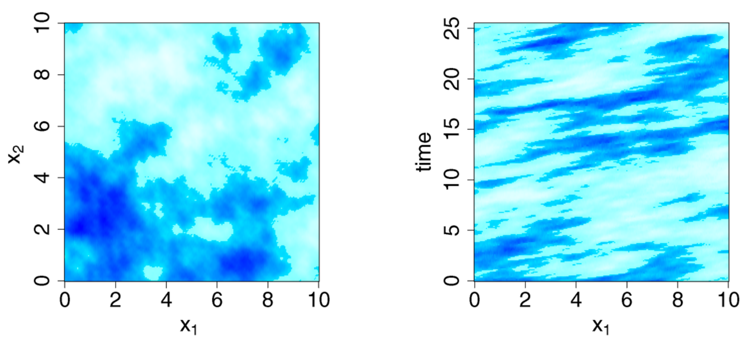

Figure 1. Realizations of the Cox–Isham covariance model in $\mathbb{R}^2 \times \mathbb{R}$. Left time $t = 0$, right $x_2 = 0$. See Example 9 for details.

function where $\varphi = W_1$ is the Whittle–Matérn model, $\mu = (1, 1)$ and

$$D = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

Remark 10. Stein (2005b) proposes models in $\mathbb{R}^d$ of the form

$$C(x, y) = \frac{\varphi([(x-y)^\top (f(x)+f(y))^{-1}(x-y)]^{1/2})}{\sqrt{|f(x)+f(y)|}}, \qquad x, y \in \mathbb{R}^d,$$

in which the values of $f : \mathbb{R}^{2d} \to \mathbb{R}^{m\times m}$ are strictly positive definite matrices, see also Paciorek (2003) and Porcu *et al.* (2009). Here, $f(x) + f(y)$ is not a variogram in general, but the proof of Theorem 5 is still applicable if $\hat{C}$ in Equation (9) is replaced by

$$\hat{C}(\omega, x, y) = \exp\bigl(-\omega^\top \bigl(f(x)+f(y)\bigr)\omega\bigr),$$

which is a positive definite function for all $\omega \in \mathbb{R}^m$.

Remark 11. The covariance model (7), which is valid in $\mathbb{R}^d$, does not allow for negative values, hence its value is limited in some applications (Gregori *et al.* (2008)). To overcome this limitation, Ma (2005a) considers differences of positive definite functions. Let $B_1, B_2, M_1, M_2 \in \mathbb{R}^{d\times d}$ be strictly positive definite matrices. Proposition 4 yields that

$$\begin{aligned} C(h, x, y) = {} & \frac{\exp(-[h^\top (M_1 + (x-y)^\top B_1 (x-y)\mathbf{1}_{d\times d})^{-1} h])}{\sqrt{|M_1 + (x-y)^\top B_1 (x-y)\mathbf{1}_{d\times d}|}} \\ & + b\frac{\exp(-[h^\top (M_2 + (x-y)^\top B_2 (x-y)\mathbf{1}_{d\times d})^{-1} h])}{\sqrt{|M_2 + (x-y)^\top B_2 (x-y)\mathbf{1}_{d\times d}|}}, \qquad h, x, y \in \mathbb{R}^d, \end{aligned}$$

is a positive definite function in $\mathbb{R}^{2d}$ that is translation invariant in its first argument if and only if for all $\omega \in \mathbb{R}^d$,

$$\begin{aligned} \hat{C}_\omega(x, y) = {} & \exp\bigl(-\omega^\top M_1 \omega - \|\omega\|^2 (x-y)^\top B_1 (x-y)\bigr) \\ & + b\exp\bigl(-\omega^\top M_2 \omega - \|\omega\|^2 (x-y)^\top B_2 (x-y)\bigr), \qquad x, y \in \mathbb{R}^d, \end{aligned}$$

is a positive definite function, that is, if and only if for all $\omega, \xi \in \mathbb{R}^d$,

$$|B_1|^{-1/2} \exp(-\omega^\top M_1 \omega - \|\omega\|^2 \xi^\top B_1^{-1} \xi) + b|B_2|^{-1/2} \exp(-\omega^\top M_2 \omega - \|\omega\|^2 \xi^\top B_2^{-1} \xi) \geq 0.$$

This is true for some negative value of b if and only if both $M_2 - M_1$ and $B_2^{-1} - B_1^{-1}$ are positive definite matrices. In this case, $C(h, x, y) : \mathbb{R}^{3d} \to \mathbb{R}$ is a positive definite function in $\mathbb{R}^{2d}$ if and only if

$$b \geq -\sqrt{|B_2|/|B_1|}.$$

Then, C_0 given by $C_0(x, y) = C(x-y, x, y)$ is a stationary covariance function in $\mathbb{R}^d$ that may take negative values.

Remark 12. The condition that $M + G(x, y)$ is strictly positive definite for all $x, y \in \mathbb{R}^d$ can be relaxed. For example, let $d = 2$ and $(h, u) = x - y \in \mathbb{R}^2$. Then, the function $C(h, u) = |u|^{-1/2} \exp(-h^2/|u|)$ is of the form (7) and defines a covariance function of a stationary, generalized random field on $\mathbb{R}^2$, see Chapter 3 in Gel'fand and Vilenkin (1964) and Chapter 17 in Koralov and Sinai (2007). Note that, here, $\lim_{u\to 0} C(0, u) = \infty$. Hence, C cannot be a translation invariant covariance function in the usual sense.

4. Model constructions based on dependent processes

The idea of the subsequent two constructions is based on the following observation. Let $C(h, u) = C_0(h)C_1(u)$, $h \in \mathbb{R}^d$, $u \in \mathbb{R}$, be a translation invariant, real-valued covariance model in $\mathbb{R}^{d+1}$ and assume we are interested in the corresponding random field at some fixed locations $x_1, \dots, x_n \in \mathbb{R}^d$ and for all $t \in \mathbb{R}$. Let Y_x, $x \in \mathbb{R}^d$, be i.i.d. temporal processes with covariance function C_1. Then

$$Z(t) = (Z_{x_1}(t), \dots, Z_{x_n}(t)) = \bigl(C_0(x_p - x_q)\bigr)^{1/2}_{p,q=1,\dots,n} (Y_{x_1}(t), \dots, Y_{x_n}(t))^\top, \qquad t \in \mathbb{R},$$

has the required covariance structure. Now, Z can be interpreted as a finite, weighted sum over Y_x, $x \in \mathbb{R}^d$. The separability is caused by the fact that Y enters into the sum only through the fixed instance t. Non-separable models can be obtained if the argument of Y also depends on the location.

4.1. Moving averages based on fields of temporal processes

Assume that $Y(A, t)$, $A \in \mathcal{B}^d$ and $t \in \mathbb{R}^l$, is a stationary process such that $Y(A_1, \cdot), \dots, Y(A_n, \cdot)$ are independent for any disjoint sets $A_1, \dots, A_n \in \mathcal{B}^d$, $n \in \mathbb{N}$. In the second argument, Y is a stationary, zero mean Gaussian random field on $\mathbb{R}^l$ with covariance function $|A|C_1$, $C_1 : \mathbb{R}^l \to \mathbb{R}$. Then,

$$\operatorname{Cov}(Y(A, t), Y(B, s)) = |A \cap B| C_1(t - s)$$

for any $s, t \in \mathbb{R}^l$ and $A, B \in \mathcal{B}^d$. Let $f : \mathbb{R}^d \to \mathbb{R}^l$ be continuous, $g : \mathbb{R}^d \to \mathbb{R}$ be continuous and square-integrable, and

$$Z(x, t) = \int_{\mathbb{R}^d} g(v - x) Y\bigl(\mathrm{d}v, f(v - x) - t\bigr), \qquad x \in \mathbb{R}^d, t \in \mathbb{R}^l.$$

Then Z is weakly stationary on $\mathbb{R}^{d+l}$ with translation invariant covariance function

$$C(h, u) = \int_{\mathbb{R}^d} g(v)g(v + h)C_1\bigl(f(v) - f(v + h) - u\bigr)\,\mathrm{d}v, \qquad h \in \mathbb{R}^d, u \in \mathbb{R}^l.$$

Example 13. Let $g(v) = (2\pi^{-1})^{d/4} \exp(-\|v\|^2)$, $v \in \mathbb{R}^d$, $l = 1$, $C_1(u) = \exp(-u^2)$, $u \in \mathbb{R}$, and $f(v) = v^\top A v + z^\top v$, $v \in \mathbb{R}^d$, for a symmetric, not necessarily positive definite matrix $A \in \mathbb{R}^{d\times d}$

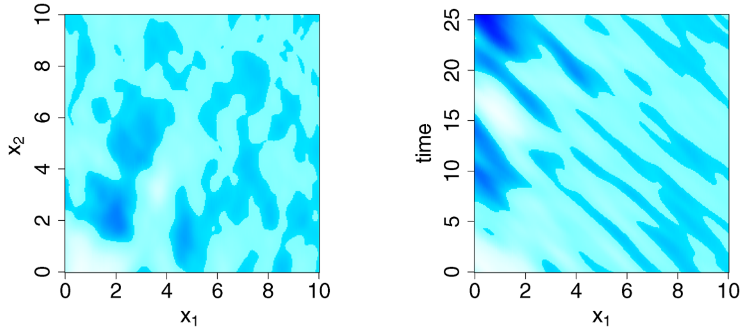

Figure 2. Realizations of a moving average random field in $\mathbb{R}^2 \times \mathbb{R}$. Left time $t = 0$, right $x_2 = 0$. See Example 13 for the definition of the covariance structure.

and $z \in \mathbb{R}^d$. Let us further introduce a non-negative random scale V, that is,

$$Z(x,t) = V^{d/2} \int_{\mathbb{R}^d} g\big(\sqrt{V}(v - x)\big) Y\big(\mathrm{d}v, \sqrt{V}\big(f(v-x) - t\big)\big), \qquad x \in \mathbb{R}^d, t \in \mathbb{R}.$$

Let $B = Ahh^\top A$. Then the covariance function of Z equals

$$C(h,u) = |\mathbf{1}_{d\times d} + 2B|^{-1/2} \mathbb{E}_V \mathrm{e}^{-V[\|h\|^2/2 + (z^\top h + u)^2 (1 - 2h^\top A(\mathbf{1}_{d\times d} + 2B)^{-1} Ah)]}, \tag{11}$$

please refer to the appendix for a proof. Equation (11) reveals that C is a potential covariance model for rainfall with frozen wind direction. Figure 2 depicts realizations of a random field with the above covariance function where $\mathbb{E}_V \exp(-VQ)$ is the Whittle–Matérn model $W_1(\sqrt{Q})$, $Q \ge 0$, $z = (2,0)$ and $A = \begin{pmatrix} 0.5 & 0 \\ 0 & 1 \end{pmatrix}$.

4.2. Models based on a single temporal process

Another class of models may be obtained by considering only a single process Y. Although the subsequent approach might be generalized, an explicit model has currently only been found within the framework of normal scale mixtures. For $x \in \mathbb{R}^d$ let

$$Z(x) = (2V/\pi)^{d/4} |S_x|^{1/4} \mathrm{e}^{-V(U-x)^\top S_x (U-x)} Y\big(\sqrt{V}\big(\xi_1(U-x) + \xi_2(x)\big)\big) \frac{g(V,x)}{\sqrt{f(U)}}. \tag{12}$$

Here, V is a positive random variable and U is a d-dimensional random variable with strictly positive density f. The one-dimensional random process Y is assumed to be stationary with Gaussian covariance function $C(t) = \mathrm{e}^{-t^2}$. The matrix S_x is strictly positive definite for all $x \in \mathbb{R}^d$, $\xi_2 : \mathbb{R}^d \to \mathbb{R}$ is arbitrary, and g is a positive function such that $\mathbb{E}_V g(V,x)^2$ is finite for all $x \in \mathbb{R}^d$. The function ξ_1 is quadratic, that is,

$$\xi_1(x) = x^\top M x + z^\top x$$

for a symmetric $d \times d$ matrix M and an arbitrary vector $z \in \mathbb{R}^d$. Let

$$c = -z^\top (x-y) + \xi_2(x) - \xi_2(y),$$
$$A = S_x + S_y + 4M(x-y)(x-y)^\top M,$$
$$m = (x-y)^\top M(x-y),$$

and

$$Q(x,y) = c^2 - m^2 + (x-y)^\top \bigl(S_x + 2(m+c)M\bigr) A^{-1} \bigl(S_y + 2(m-c)M\bigr)(x-y).$$

Then the covariance function of Z equals

$$C(x,y) = \frac{2^{d/2}|S_x|^{1/4}|S_y|^{1/4}}{\sqrt{|A|}} \cdot \mathbb{E}_V g(V,x) g(V,y) \exp(-VQ(x,y)), \qquad x,y \in \mathbb{R}^d. \qquad (13)$$

The proof is given in the Appendix.

Example 14. Translation-invariant models in $\mathbb{R}^d$ are obtained if both S_x and g do not depend on x. Assume S_x is twice the identity matrix, $g(v) = (2\sqrt{v})^{1-\nu}/\sqrt{\Gamma(\nu)}$, $v, \nu > 0$, and V follows the Fréchet distribution $F(v) = \mathrm{e}^{-1/(4v)}$, $v > 0$. Two particular models might be of special interest, either because of their simplicity or their explicit spatio-temporal modelling. First, if $c \equiv 0$ then

$$C(h) = \frac{W_\nu(\|h\|)}{|\mathbf{1}_{d\times d} + Mhh^\top M|^{1/2}}, \qquad h \in \mathbb{R}^d,$$

according to formula 3.471.9 in Gradshteyn and Ryzhik (2000). Second, an explicit spatio-temporal model in $\mathbb{R}^{d+1}$ is obtained for

$$\xi_2(x,t) = t, \qquad x \in \mathbb{R}^d, t \in \mathbb{R}, \quad \text{and} \quad M = \begin{pmatrix} L & 0 \\ 0 & 0 \end{pmatrix}.$$

Then, with $D = \mathbf{1}_{d\times d} + Lhh^\top L$, we get

$$C(h,u) = |D|^{-1/2} W_\nu\bigl(\sqrt{Q(h,u)}\bigr), \qquad h \in \mathbb{R}^d, u \in \mathbb{R},$$

where

$$Q(h,u) = (u - z^\top h)^2 - (h^\top Lh)^2 + h^\top \bigl(D + (u - z^\top h)L\bigr) D^{-1} \bigl(D + (u - z^\top h)L\bigr) h.$$

Example 15. Let $\xi_1 \equiv \xi_2 \equiv 0$. Then the random process $Y(t)$ is considered only at instance $t = 0$ and the exponent $Q(x,y)$ simplifies to

$$Q(x,y) = (x-y)^\top S_x (S_x + S_y)^{-1} S_y (x-y) = (x-y)^\top (S_x^{-1} + S_y^{-1})^{-1}(x-y).$$

Let $g(v,x) = (2\sqrt{v})^{1-\nu(x)}/\Gamma(\nu(x))^{1/2}$, ν a positive function on $\mathbb{R}^d$, and V a Fréchet variable with distribution function $F(v) = \mathrm{e}^{-1/(4v)}$, $v > 0$. Then, the first model given in Stein (2005b) is obtained,

$$C(x,y) = \frac{2^{d/2}|S_x|^{1/4}|S_y|^{1/4}\Gamma((\nu(x)+\nu(y))/2)}{[|S_x+S_y|\Gamma(\nu(x))\Gamma(\nu(y))]^{1/2}} W_{(\nu(x)+\nu(y))/2}(Q(x,y)^{1/2}), \qquad x,y \in \mathbb{R}^d.$$

The second model given in Stein (2005b), a generalization of the Cauchy model, is obtained by $g(v,x) = v^{(\delta(x)-1)/2}$ and a standard exponential random variable V, that is,

$$C(x,y) = \frac{2^{d/2}|S_x|^{1/4}|S_y|^{1/4}}{|S_x+S_y|^{1/2}(1+Q(x,y))^{(\delta(x)+\delta(y))/2}}, \qquad x,y \in \mathbb{R}^d.$$

If ν and δ are constant, then the above models are special cases of Theorem 5.

See Theorem 1 in Porcu *et al.* (2009) for a class of models that generalizes Stein's examples.

Example 16. A cyclone can be mimicked if rotation matrices are included in the model,

$$C(x,y) = \frac{2^{d/2}|S_x|^{1/4}|S_y|^{1/4}}{\sqrt{|S_x+S_y|}} W_\nu\big(\big(h^\top S_x(S_x+S_y)^{-1}S_y h\big)^{1/2}\big), \qquad x,y, \in \mathbb{R}^3,$$

where

$$S_x = \mathrm{diag}(1,1,1) + R(x)^\top A^\top x x^\top A R(x), \qquad A \in \mathbb{R}^{3\times 3},$$

$$R(x) = \begin{pmatrix} \cos(\alpha x_3) & -\sin(\alpha x_3) & 0 \\ \sin(\alpha x_3) & \cos(\alpha x_3) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad x = (x_1,x_2,x_3) \in \mathbb{R}^3, \alpha \in \mathbb{R},$$

and

$$h = x^\top R(x) - y^\top R(y).$$

The positive definiteness of the model is now ensured by both Theorem 5 and a generalized version of Z in Equation (12), replacing x by $x^\top R(x)$ there. Note that $x \mapsto x^\top R(x)$ is a bijection. Figure 3 depicts realizations of a random field with the above covariance function where $\alpha = -2\pi$, $\nu = 1$, and

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

5. Multivariate spatio-temporal models

Here, we generalize Theorem 5 to construct multivariate cross covariance functions. Let $\underline{M} = (M + M^\top)/2$ for any real-valued square matrix M.

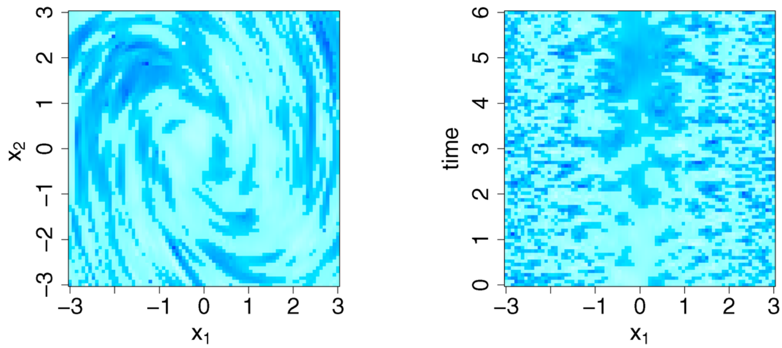

Figure 3. Realizations of a random field in $\mathbb{R}^3$ that mimics a cyclone. Left time $x_3 = 0$, right $x_2 = 0$. See Example 16 for the definition of the covariance structure.

Theorem 17. *Assume that l, m and d are positive integers, $A_j \in \mathbb{R}^{l\times d}$ for $j = 1, \dots, m$. Suppose that φ is a normal scale mixture and $G : \mathbb{R}^{2d} \to \mathbb{R}^{l\times l}$ is a cross covariance function. Let $M \in \mathbb{R}^{d\times d}$ be a positive definite matrix such that $M - A_j^\top G(x, y)A_k$ is strictly positive definite for all $x, y \in \mathbb{R}^d$ and $j, k = 1, \dots, d$. Then $C = (C_{jk})_{j,k=1,\dots,m}$ is a cross covariance function in $\mathbb{R}^d$ for*

$$C_{jk}(x, y) = \frac{\varphi([(x-y)^\top (M - A_j^\top G(x, y)A_k)^{-1}(x-y)]^{1/2})}{\sqrt{|M - A_j^\top G(x, y)A_k|}}, \quad x, y \in \mathbb{R}^d,\ j, k = 1, \dots, m. \tag{14}$$

Proof. Lemma 6 yields that

$$\begin{aligned}(\omega^\top A_j^\top G(x, y)A_k\omega)_{j,k=1,\dots,m} &= (\omega^\top A_j^\top G(x, y)A_k\omega)_{j,k=1,\dots,m} \\ &= (A_1\omega, \dots, A_m\omega)^\top G(x, y)(A_1\omega, \dots, A_m\omega)\end{aligned}$$

is a cross covariance function for all $\omega \in \mathbb{R}^d$. Part 1 of Theorem 1 yields that $C_\omega(x, y) = (\exp(\omega^\top A_j^\top G(x, y)A_k\omega))_{j,k=1,\dots,m}$ is also a cross covariance function. By assumption, $M - A_j^\top G(x, y)A_k$ is strictly positive definite. Hence, as a result of Proposition 4, the Fourier transform of the function $\omega \mapsto \exp(-\omega^\top M\omega)C_\omega(x, y)$ is a cross covariance function, which is of the form (14). □

Appendix

A.1. Proof for the covariance function in Example 9

Let $f_{\mu,D/2}(x)$ be the multivariate normal density with expectation μ and covariance matrix $D/2$. Then we get

$$\begin{aligned}
&-\log\bigl(\varphi(h-uv)f_{\mu,D/2}(v)\bigr)+\tfrac{1}{2}\log((2\pi)^d|D|)\\
&\quad=h^\top h-2uh^\top v+u^2v^\top v+v^\top D^{-1}v-2\mu^\top D^{-1}v+\mu^\top D^{-1}\mu\\
&\quad=h^\top h+\mu^\top D^{-1}\mu+(v-\xi)^\top(u^2\mathbf{1}_{d\times d}+D^{-1})(v-\xi)-\xi^\top(u^2\mathbf{1}_{d\times d}+D^{-1})\xi
\end{aligned}$$

with $\xi=(u^2\mathbf{1}_{d\times d}+D^{-1})^{-1}(uh+D^{-1}\mu)$. Hence,

$$\begin{aligned}
&-\log C(h,u)+\tfrac{1}{2}\log(|D|)+\tfrac{1}{2}\log(|u^2\mathbf{1}_{d\times d}+D^{-1}|)\\
&\quad=h^\top h+\mu^\top D^{-1}\mu-\xi^\top(u^2\mathbf{1}_{d\times d}+D^{-1})\xi\\
&\quad=(h-u\mu)^\top(\mathbf{1}_{d\times d}+u^2D)^{-1}(h-u\mu)
\end{aligned}$$

which yields the assertion.

A.2. Proof for the covariance function in Example 13

We proof the formula for the covariance function in Example 13, but also demonstrate that a slightly more general function g does not give a more general model. To this end, let $g(v)=(|2\pi^{-1}M|)^{1/4}\exp(-v^\top Mv)$, $v\in\mathbb{R}^d$, for a strictly positive definite matrix $M\in\mathbb{R}^{d\times d}$. For ease of notation we assume that $V\equiv 1$. Then

$$\begin{aligned}
&-\log\bigl(g(v)g(v+h)C_1\bigl(f(v)-f(v+h)-u\bigr)\bigr)-\tfrac{1}{2}\log(|2\pi^{-1}M|)\\
&\quad=v^\top Mv+(v+h)^\top M(v+h)+(2v^\top Ah+h^\top Ah+z^\top h+u)^2\\
&\quad=2v^\top Mv+4v^\top Bv+2v^\top(2B+M+2uA+2Ahz^\top)h+c
\end{aligned}$$

where $B=Ahh^\top A$ and $c=[h^\top Ah+z^\top h+u]^2+h^\top Mh$. Hence, with $D=2B+M+2[u+z^\top h]A$,

$$\begin{aligned}
&-\log\bigl(g(v)g(v+h)C_1\bigl(f(v)-f(v+h)+u\bigr)\bigr)-\tfrac{1}{2}\log(|2\pi^{-1}M|)\\
&\quad=\bigl(v-(2M+4B)^{-1}Dh\bigr)^\top(2M+4B)\bigl(v-(2M+4B)^{-1}Dh\bigr)\\
&\qquad-h^\top D(2M+4B)^{-1}Dh+c.
\end{aligned}$$

Thus,

$$C(h,u)=\frac{|M|^{1/2}}{|M+2B|^{1/2}}\exp\bigl(-c+h^\top D(2M+4B)^{-1}Dh\bigr),\qquad h\in\mathbb{R}^d,u\in\mathbb{R}.$$

Let $M^{-1/2}$ be a symmetric matrix with $M^{-1/2}MM^{-1/2} = \mathbf{1}_{d\times d}$. Replacing on the right hand side $M^{-1/2}AM^{-1/2}$ by $\tilde{A}$, $M^{-1/2}z$ by $\tilde{z}$ and $M^{1/2}h$ by $\tilde{h}$ shows that M causes nothing but a geometrical anisotropy effect. Hence, we may assume that M is the identity matrix. Then

$$C(h,u) = |\mathbf{1}_{d\times d} + 2B|^{-1/2} \exp\bigl(-\bigl[c - \tfrac{1}{2}h^\top D(\mathbf{1}_{d\times d} + 2B)^{-1}Dh\bigr]\bigr)$$

which yields Equation (11).

A.3. Proof of Equation (13)

Let $h = x - y$ and $w = U - x$. Then we have

$$\begin{aligned}\operatorname{Cov}(Z(x), Z(y)) = {}& \pi^{-d/2}|S_x|^{1/4}|S_y|^{1/4}\mathbb{E}_V V^{d/2} g(V,x)g(V,y)\\ &\times \int \exp\bigl(-Vw^\top S_x w - V(w+h)^\top S_y(w+h)\\ &\qquad - V\bigl(w^\top Mw - (w+h)^\top M(w+h) + c\bigr)^2\bigr)\,\mathrm{d}w.\end{aligned}$$

The value of the integral is at most $\int \exp(-Vw^\top S_x w)\,\mathrm{d}w$. Hence $\operatorname{Cov}(Z(x), Z(y)) < \infty$ if $\mathbb{E}_V g(V,x)g(V,y) < \infty$. Now,

$$\begin{aligned}&w^\top S_x w + (w+h)^\top S_y(w+h) + \bigl(w^\top Mw - (w+h)^\top M(w+h) + c\bigr)^2\\ &\quad = w^\top(S_x + S_y + 4Mhh^\top M)w + 2w^\top\bigl(S_y + 2(h^\top Mh - c)M\bigr)h + h^\top S_y h + (h^\top Mh - c)^2\\ &\quad = (w-\mu)^\top A(w-\mu) - \mu^\top A\mu + h^\top S_y h + (h^\top Mh - c)^2\end{aligned}$$

with $\mu = -A^{-1}(S_y + 2(h^\top Mh - c)M)h$. That is,

$$\begin{aligned}\operatorname{Cov}(Z(x), Z(y)) = {}& |A|^{-1/2}|S_x|^{1/4}|S_y|^{1/4}\mathbb{E}_V g(V,x)g(V,y)\\ &\times \mathrm{e}^{-V[hS_y h + (h^\top Mh - c)^2 - \mu^\top A\mu]}.\end{aligned} \tag{15}$$

On the other hand, using the transform $w = U - y$, we get

$$\begin{aligned}&\operatorname{Cov}(Z(x), Z(y))\\ &\quad = \pi^{-d/2}|S_x|^{1/4}|S_y|^{1/4}\mathbb{E}_V V^{d/2} g(V,x)g(V,y)\\ &\qquad \times \int \exp\bigl(-V(w-h)^\top S_x(w-h) + -VhS_y h\\ &\qquad\qquad - V\bigl((w-h)^\top M(w-h) - w^\top Mw + c\bigr)^2\bigr)\,\mathrm{d}w\\ &\quad = |A|^{-1/2}|S_x|^{1/4}|S_y|^{1/4}\mathbb{E}_V g(V,x)g(V,y)\mathrm{e}^{-V[hS_x h + (h^\top Mh + c)^2 - \nu^\top A\nu]}\end{aligned} \tag{16}$$

with $\nu = A^{-1}(S_x + 2(h^\top Mh + c)M)h$.

Choosing $V \equiv 1$ and g a constant function we obtain that the exponents in (15) and (16) must be equal, that is,

$$\begin{aligned}
&hS_yh + (h^\top Mh - c)^2 - \mu^\top A\mu \\
&\quad = \tfrac{1}{2}[hS_yh + (h^\top Mh - c)^2 - \mu^\top A\mu + hS_xh + (h^\top Mh + c)^2 - \nu^\top A\nu] \\
&\quad = \tfrac{1}{2}[h(S_y + S_x + 4Mhh^\top M)h - 2(h^\top Mh)^2 + 2c^2 - (\mu - \nu)A(\mu - \nu) - 2\nu^\top A^{-1}\mu] \\
&\quad = c^2 - (h^\top Mh)^2 - \nu^\top A^{-1}\mu.
\end{aligned}$$

Acknowledgements

The author is grateful to Zakhar Kabluchko, Emilio Porcu and the referees for valuable suggestions and comments.

References

Barndorff-Nielsen, O. (1979). Hyperbolic distributions and distributions on hyperbolae. *Scand. J. Statist.* **5** 151–157. MR0509451

Berg, C., Christensen, J.P.R. and Ressel, P. (1984). *Harmonic Analysis on Semigroups. Theory of Positive Definite and Related Functions*. New York: Springer. MR0747302

Berlinet, A. and Thomas-Agnan, C. (2004). *Reproducing Kernel Hilbert Spaces in Probability and Statistics*. Boston: Kluwer.

Cox, D.R. and Isham, V.S. (1988). A simple spatial-temporal model of rainfall. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **415** 317–328. MR0932924

Cressie, N.A.C. and Huang, H.-C. (1999). Classes of nonseparable, spatio-temporal stationary covariance functions. *J. Amer. Statist. Assoc.* **94** 1330–1340. MR1731494

Gel'fand, I.M. and Vilenkin, N.Y. (1964). *Generalized Functions: Applications of Harmonic Analysis*. New York: Academic Press.

Gneiting, T. (1997). Normal scale mixtures and dual probability densities. *J. Stat. Comput. Simul.* **59** 375–384.

Gneiting, T. (2002). Nonseparable, stationary covariance functions for space-time data. *J. Amer. Statist. Assoc.* **97** 590–600. MR1941475

Gneiting, T. and Schlather, M. (2004). Stochastic models that separate fractal dimension and the Hurst effect. *SIAM Rev.* **46** 269–282. MR2114455

Gneiting, T., Sasvári, Z. and Schlather, M. (2001). Analogies and correspondences between variograms and covariance functions. *Adv. in Appl. Probab.* **33** 617–630. MR1860092

Gneiting, T., Genton, M.G. and Guttorp, P. (2007). Geostatistical space-time models, stationarity, separability and full symmetry. In *Statistical Methods for Spatio-Temporal Systems* (B. Finkenstadt, L. Held, and V. Isham, eds.) chap. 4. Boca Raton: Chapman & Hall/CRC.

Gradshteyn, I.S. and Ryzhik, I.M. (2000). *Table of Integrals, Series, and Products*, 6th ed. London: Academic Press. MR1773820

Gregori, P., Porcu, E., Mateu, J. and Sasvári, Z. (2008). On potentially negative space time covariances obtained as sum of products of marginal ones. *Ann. Inst. Statist. Math.* **60** 865–882. MR2453575

Kolovos, A., Christakos, G., Hristopulos, D.T. and Serre, M.L. (2004). Methods for generating non-separable spatiotemporal covariance models with potential environmental applications. *Adv. Water. Res.* **27** 815–830.

Koralov, L.B. and Sinai, Y.G. (2007). *Theory of Probability and Random Processes*, 2nd ed. Berlin: Springer. MR2343262

Ma, C. (2002). Spatio-temporal covariance functions generated by mixtures. *Math. Geol.* **34** 965–975. MR1951438

Ma, C. (2003). Families of spatio-temporal stationary covariance models. *J. Stat. Plann. Inference* **116** 489–501. MR2000096

Ma, C. (2005a). Semiparametric spatio-temporal covariance models with the arma temporal margin. *Ann. Inst. Statist. Math.* **57** 221–233. MR2160648

Ma, C. (2005b). Spatio-temporal variograms and covariance models. *Adv. in Appl. Probab.* **37** 706–725. MR2156556

Ma, C. (2005c). Linear combinations of space-time covariance functions and variograms. *IEEE Trans. Signal Process.* **53** 857–864. MR2123904

Ma, C. (2008). Recent developments on the construction of spatio-temporal covariance models. *Stoch. Environ. Res. Risk. Asses.* **22** 39–47. MR2418410

Mateu, J., Porcu, E. and Gregori, P. (2008). Recent advances to model anisotropic space-time data. *Stat. Methods Appl.* **17** 209–223. MR2395119

Matheron, G. (1972). Leçon sur les fonctions aléatoire d'ordre 2. Technical Report C-53, Ecole des Mines de Paris. Available at http://cg.ensmp.fr/bibliotheque/public/MATHERON_Cours_00302.pdf.

Paciorek, C. (2003). Nonstationary Gaussian processes for regression and spatial modelling. Ph.D. thesis, Carnegie Mellon Univ., Dept. Statistics. Available at www.biostat.harvard.edu/∼paciorek/diss/paciorek-thesis.pdf.

Porcu, E., Mateu, J. and Christakos, G. (2009). Quasi-arithmetic means of covariance functions with potential applications to space-time data. *J. Multivariate Anal.* **100** 1830–1844. MR2535390

Reisert, M. and Burkhardt, H. (2007). Learning equivariant functions with matrix valued kernels. *J. Mach. Learn. Res.* **8** 385–408. MR2320676

Schoenberg, I.J. (1938a). Metric spaces and completely monotone functions. *Ann. Math.* **39** 811–841. MR1503439

Schoenberg, I.J. (1938b). Metric spaces and positive definite functions. *Trans. Amer. Math. Soc.* **44** 522–536. MR1501980

Stein, M.L. (1999). *Interpolation of Spatial Data*. Heidelberg, New York: Springer. MR1697409

Stein, M.L. (2005a). Space-time covariance functions. *J. Amer. Statist. Assoc.* **100** 310–321. MR2156840

Stein, M.L. (2005b). Nonstationary spatial covariance functions. Technical report, Univ. Chicago. Available at www.stat.uchicago.edu/cises/research/cises-tr21.pdf.

Stein, M.L. (2005c). Statistical methods for regular monitoring data. *J. Roy. Statist. Soc. Ser. B* **67** 667–687. MR2210686

Wackernagel, H. (2003). *Multivariate Geostatistics*, 3rd ed. Berlin, Heidelberg: Springer.

Wendland, H. (2005). *Scattered Data Approximation*. Cambridge: Cambridge Univ. Press. MR2131724

Yaglom, A.M. (1987a). *Correlation Theory of Stationary and Related Random Functions I, Basic Results*. New York, Berlin: Springer. MR0893393

Yaglom, A.M. (1987b). *Correlation Theory of Stationary and Related Random Functions II, Supplementary Notes and References*. New York, Berlin: Springer. MR0915557

Received November 2008 and revised May 2009