Biometrika (2013), **100**, 1, *pp.* 213–220
doi: 10.1093/biomet/ass057
Advance Access publication 29 November 2012

Miscellanea

Spatially varying cross-correlation coefficients in the presence of nugget effects

By WILLIAM KLEIBER

Department of Applied Mathematics, University of Colorado, Boulder, Colorado 80309, U.S.A.
william.kleiber@colorado.edu

AND MARC G. GENTON

CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia
marc.genton@kaust.edu.sa

Summary

We derive sufficient conditions for the cross-correlation coefficient of a multivariate spatial process to vary with location when the spatial model is augmented with nugget effects. The derived class is valid for any choice of covariance functions, and yields substantial flexibility between multiple processes. The key is to identify the cross-correlation coefficient matrix with a contraction matrix, which can be either diagonal, implying a parsimonious formulation, or a fully general contraction matrix, yielding greater flexibility but added model complexity. We illustrate the approach with a bivariate minimum and maximum temperature dataset in Colorado, allowing the two variables to be positively correlated at low elevations and nearly independent at high elevations, while still yielding a positive definite covariance matrix.

Some key words: Cross-correlation; Multivariate process; Nonstationarity; Random field; Spatial statistic.

1. Introduction

Consider a multivariate process $Z(s) = \{Z_1(s), \ldots, Z_p(s)\}^{\mathrm{T}}$ indexed by locations $s \in \mathbb{R}^d$ with matrix-valued covariance function

$$C(s_1, s_2) = \begin{pmatrix} \mathrm{cov}\{Z_1(s_1), Z_1(s_2)\} & \cdots & \mathrm{cov}\{Z_1(s_1), Z_p(s_2)\} \\ \vdots & \ddots & \vdots \\ \mathrm{cov}\{Z_p(s_1), Z_1(s_2)\} & \cdots & \mathrm{cov}\{Z_p(s_1), Z_p(s_2)\} \end{pmatrix}. \tag{1}$$

We call the matrix-valued function (1) nonnegative definite if the covariance matrix of $Z = \{Z(s_1)^{\mathrm{T}}, \ldots, Z(s_n)^{\mathrm{T}}\}^{\mathrm{T}}$ is nonnegative definite for any choices of $s_1, \ldots, s_n$. Let $C_{ii}(s_1, s_2) = \mathrm{cov}\{Z_i(s_1), Z_i(s_2)\}$ be the direct covariance functions, and let $\beta_{ij} C_{ij}(s_1, s_2) = \mathrm{cov}\{Z_i(s_1), Z_j(s_2)\}$ be the cross-covariance functions. The parameters β_{ij} define the co-located cross-correlation coefficients, that is, $\beta_{ij} = \mathrm{cor}\{Z_i(s), Z_j(s)\}$ $(i, j = 1, \ldots, p)$. In particular, for arbitrary valid choices of $C_{ij}(\cdot, \cdot)$, we require $\beta_{ii} = 1$, $|\beta_{ij}| \leqslant 1$ $(i, j = 1, \ldots, p)$, and the matrix with β_{ij} as its (i, j)th entry is symmetric and nonnegative definite.

The relationship between physical processes often evolves across space, so that

$$\mathrm{cor}\{Z_i(s_1), Z_j(s_1)\} \neq \mathrm{cor}\{Z_i(s_2), Z_j(s_2)\}.$$

This suggests generalizing β_{ij} so that it is a function of location, that is, $\beta_{ij} \to \beta_{ij}(s_1, s_2)$. The necessary conditions for the spatially varying cross-correlation coefficients $\{\beta_{ij}(\cdot,\cdot)\}_{i,j=1}^{p}$ to be valid depend on the class of covariance functions in use. We introduce sufficient conditions for $\{\beta_{ij}(\cdot,\cdot)\}_{i,j=1}^{p}$ to yield a matrix-valued covariance function when used with any valid covariance structure.

As an example of one possible construction of matrix-valued covariance functions, consider the multivariate Matérn model, introduced by Gneiting et al. (2010). They supposed that $C_{ij}(\cdot,\cdot)$ are Matérn covariance functions of the form

$$\mathrm{cov}\{Z_i(s_1), Z_j(s_2)\} = \beta_{ij} A(\nu_i, \nu_j)\sigma_i\sigma_j \frac{2^{1-\nu_{ij}}}{\Gamma(\nu_{ij})} (a\|s_1 - s_2\|)^{\nu_{ij}} \mathrm{K}_{\nu_{ij}}(a\|s_1 - s_2\|),$$

where $\mathrm{K}_{\nu_{ij}}(\cdot)$ is a modified Bessel function of the second kind of order $\nu_{ij} = (\nu_i + \nu_j)/2$. Here, $A(\nu_i, \nu_j)$ is a constant that depends on the marginal smoothness parameters ν_i and ν_j. Gneiting et al. (2010) derived the necessary and sufficient conditions for the scale parameter a to be process dependent, that is, $a \to a_i$, for $p = 2$, and Apanasovich et al. (2012) described sufficient conditions for arbitrary p. Kleiber & Nychka (2012) extended the multivariate Matérn model for nonstationary multivariate processes. Other multivariate random field models include latent dimensional constructions (Apanasovich & Genton, 2010; Porcu & Zastavnyi, 2011), the linear model of coregionalization (Goulard & Voltz, 1992; Gelfand et al., 2004; Schmidt & Gelfand, 2003; Wackernagel, 2003; Zhang, 2007), covariance convolution (Gaspari & Cohn, 1999; Majumdar & Gelfand, 2007; Majumdar et al., 2010), or kernel convolution (Ver Hoef & Barry, 1998); see Fanshawe & Diggle (2012) for discussion. These models are then often used for cokriging; see Furrer & Genton (2011) and references therein.

2. Characterizing the cross-correlation coefficient

2·1. *Set-up*

For the remainder of this article, assume that the statistical model augments the spatially correlated process $Z(s)$ with an additive nugget effect $\varepsilon(s) = \{\varepsilon_1(s), \ldots, \varepsilon_p(s)\}^{\mathrm{T}}$, a zero-mean spatial white noise process with variance matrix $\mathrm{diag}\{\tau_1(s)^2, \ldots, \tau_p(s)^2\}$. The nugget effect is usually included to account for microscale variability that cannot be distinguished from measurement error. Its inclusion is critical to define our class of sufficient conditions for nonstationary cross-correlation coefficients. The necessary conditions for spatially varying $\beta_{ij}(\cdot,\cdot)$ are dependent on the exact class of multivariate covariance functions in use. This distinction is crucial; our sufficient conditions hold for any covariance class, but the necessary conditions for a specific covariance class may be different. Define the quantity

$$\gamma_i(s) = \frac{C_{ii}(s,s) + \tau_i(s)^2}{C_{ii}(s,s)}.$$

Then the full covariance matrix for the ith variable at locations $s_1, \ldots, s_n$ can be written as a Hadamard product,

$$\begin{pmatrix} \gamma_i(s_1) & 1 & \cdots & 1 \\ 1 & \gamma_i(s_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 1 & \cdots & \cdots & \gamma_i(s_n) \end{pmatrix} \odot \begin{pmatrix} C_{ii}(s_1,s_1) & C_{ii}(s_1,s_2) & \cdots & C_{ii}(s_1,s_n) \\ C_{ii}(s_2,s_1) & C_{ii}(s_2,s_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ C_{ii}(s_n,s_1) & \cdots & \cdots & C_{ii}(s_n,s_n) \end{pmatrix}$$
$$= \begin{pmatrix} C_{ii}(s_1,s_1) + \tau_i(s_1)^2 & C_{ii}(s_1,s_2) & \cdots & C_{ii}(s_1,s_n) \\ C_{ii}(s_2,s_1) & C_{ii}(s_2,s_2) + \tau_i(s_2)^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ C_{ii}(s_n,s_1) & \cdots & \cdots & C_{ii}(s_n,s_n) + \tau_i(s_n)^2 \end{pmatrix}. \quad (2)$$

With $\varepsilon = \{\varepsilon(s_1)^{\mathrm{T}}, \ldots, \varepsilon(s_n)^{\mathrm{T}}\}^{\mathrm{T}}$, the covariance matrix of $Z + \varepsilon$ is $\beta \odot C$. Here, β and C are block matrices with $p \times p$ large blocks, the (i,j)th of which have (k,ℓ)th entries $\beta_{ij}(s_k, s_\ell)$ and $C_{ij}(s_k, s_\ell)$, respectively. For the diagonal blocks of β, define $\beta_{ii}(s_k, s_\ell) = 1$ for $k \neq \ell$ and $\beta_{ii}(s_k, s_k) = \gamma_i(s_k)$, that is, the first matrix of (2).

If the matrix-valued covariance function $C(\cdot, \cdot)$ is nonnegative definite, a sufficient condition for $\beta \odot C$ to be valid is for β to be nonnegative definite. For any arbitrary valid choice of C, we characterize the possible functional forms for β. To keep $\beta \odot C$ symmetric, we let β be symmetric; asymmetry can be introduced by the method of Li & Zhang (2011).

2·2. *Bivariate case*

For the bivariate case, $p = 2$, we completely classify the set of possible cross-correlation matrices β, which depends on the following definitions. A matrix K is called a contraction matrix if its singular values are all bounded by unity. We call $M^{1/2}$ the square root matrix of M if $M^{1/2}M^{1/2} = M$.

Proposition 1. *If $p = 2$, then β is nonnegative definite if and only if $\beta_{12} = \beta_{11}^{1/2} K \beta_{22}^{1/2}$ for some contraction matrix K.*

The proof is in the Appendix. Clearly, the available variability of cross-correlation across the study domain is intimately linked to the total variance to marginal variance ratio, $\gamma_i(s)$. An immediate corollary of Proposition 1 is that if there is no nugget effect, i.e., $\tau_i(s) = 0$ $(i = 1, 2)$, then the only valid cross-correlation matrix β_{12} is a constant matrix, whose values are bounded by unity. For $\gamma_i(s) > 1$, there is a set of valid functions $\beta_{12}(\cdot, \cdot)$.

It is instructive to examine the possible combinations of co-located cross-correlation coefficients $\beta_{12}(s_1, s_1)$ and $\beta_{12}(s_2, s_2)$ that are valid in a bivariate process for two locations s_i $(i = 1, 2)$. Suppose that the nugget variance does not vary with location and that both processes have common stationary total-to-marginal variance ratios $\gamma_1(s) = \gamma_2(s) = \gamma$, noting that such an assumption is not necessary, but simplifies exposition. Then

$$\beta_{12} = \begin{pmatrix} \gamma & 1 \\ 1 & \gamma \end{pmatrix}^{1/2} K \begin{pmatrix} \gamma & 1 \\ 1 & \gamma \end{pmatrix}^{1/2}, \qquad \begin{pmatrix} \gamma & 1 \\ 1 & \gamma \end{pmatrix}^{1/2} = \begin{pmatrix} a & b \\ b & a \end{pmatrix},$$

where $b = [\{\gamma + (\gamma^2 - 1)^{1/2}\}/2]^{1/2}$ and $a = 1/(2b)$. It is convenient to write K as its singular value decomposition $K = USV$, where $\max\{\text{diag}(S)\} \leqslant 1$. The simplest case to use in practice lets U and V be identity matrices with diagonal entries ± 1, making K diagonal, that is, $K = \text{diag}(k_1, k_2)$, where $|k_1| \leqslant 1$ and $|k_2| \leqslant 1$. Then the possible values of co-located cross-correlation are $\beta_{12}(s_1, s_1) = k_1 a^2 + k_2 b^2$, and $\beta_{12}(s_2, s_2) = k_1 b^2 + k_2 a^2$. This defines a rhombus in $\mathbb{R}^2$, whose four vertices are formed when $(k_1, k_2) = (\pm 1, \pm 1)$.

Allowing the contraction matrix K to be a general contraction, rather than strictly diagonal, we can write, in singular value decomposition form,

$$K = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$$

for some angles θ, ϕ. Then the set of valid pairs is

$$\begin{aligned} \beta_{12}(s_1, s_1) = {} & s_1\{a^2 \cos\theta \cos\phi + ab(\cos\phi \sin\theta - \cos\theta \sin\phi) - b^2 \sin\theta \sin\phi\} \\ & + s_2\{-a^2 \sin\theta \sin\phi - ab(\cos\phi \sin\theta - \cos\theta \sin\phi) + b^2 \cos\theta \cos\phi\}, \\ \beta_{12}(s_2, s_2) = {} & s_1\{-a^2 \sin\theta \sin\phi + ab(\cos\phi \sin\theta - \cos\theta \sin\phi) + b^2 \cos\theta \cos\phi\} \\ & + s_2\{a^2 \cos\theta \cos\phi - ab(\cos\phi \sin\theta - \cos\theta \sin\phi) - b^2 \sin\theta \sin\phi\}. \end{aligned}$$

Figure 1 shows the sets of valid pairs $\{\beta_{12}(s_1, s_1), \beta_{12}(s_2, s_2)\}$ for various values of γ. When $\gamma = 1$, we recover that the cross-correlation coefficient must be spatially constant. The area of the valid regions increases as a function of γ. In Fig. 1(a), we assume K is a diagonal matrix with bounded entries, whereas in Fig. 1(b), K is a general 2×2 contraction matrix. Based on these explorations, using a general contraction matrix yields slightly larger valid regions for $\{\beta_{12}(s_1, s_1), \beta_{12}(s_2, s_2)\}$, but at the cost of added model

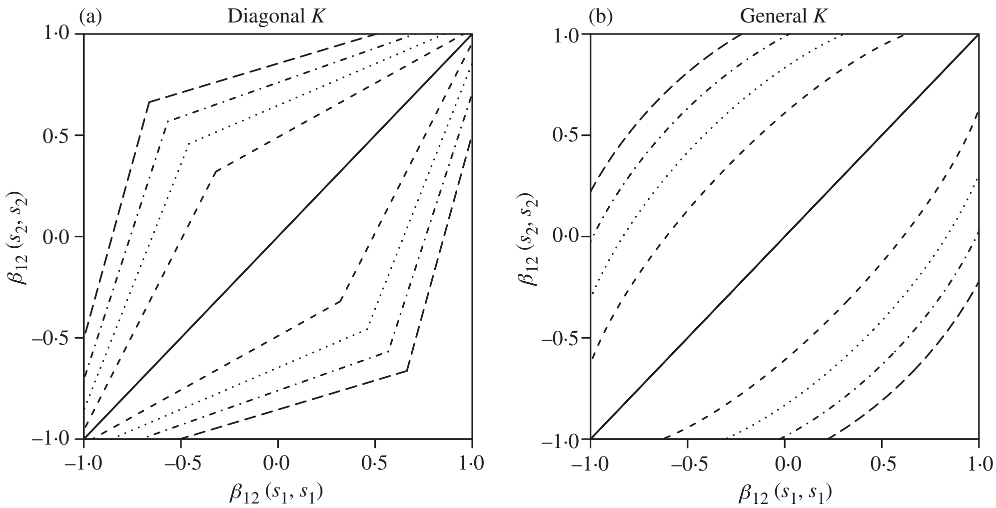

Fig. 1. Regions of valid pairs $\{\beta_{12}(s_1, s_1), \beta_{12}(s_2, s_2)\}$ for $\gamma = 1$ (solid), 1·05 (short dashed), 1·1 (dotted), 1·15 (dot dashed), and 1·2 (long dashed), assuming the contraction matrix K is either diagonal, or a general contraction matrix. For diagonal K, the valid regions are rhombi whose boundaries are shown as lines. For a general contraction K, the valid regions are ellipses whose boundaries are shown as curves.

complexity. For domains where the cross-correlation is changing slowly, or is of moderate magnitude, we recommend using a diagonal contraction matrix, as the statistical model is then more parsimonious, and still yields substantial flexibility.

2·3. *Multivariate case*

For an arbitrary number of processes, $p \geqslant 2$, we consider the multivariate generalization of a contraction matrix. The following theorem characterizes the class of possible cross-correlation coefficients for any number of processes with a valid but arbitrary covariance structure. Its proof is in the Appendix.

THEOREM 1. *Let K be an $np \times np$ block matrix with $p \times p$ large blocks, $\{K_{ij}\}_{i,j=1}^{p}$, such that $K_{ii} = I$ and $K_{ij} = \beta_{ii}^{-1/2}\beta_{ij}\beta_{jj}^{-1/2}$ for $i \neq j$. Then β is nonnegative definite if and only if K is nonnegative definite.*

When $p = 2$, the result of Theorem 1 reduces to K_{12} being a contraction matrix, which we showed in Proposition 1. For the trivariate case, $p = 3$, we have

$$X \begin{pmatrix} I & K_{12} & K_{13} \\ K_{21} & I & K_{23} \\ K_{31} & K_{32} & I \end{pmatrix} X^{\mathrm{T}}$$
$$\sim \begin{pmatrix} I & 0 & 0 \\ 0 & I - K_{21}K_{12} & 0 \\ 0 & 0 & I - K_{31}K_{13} - (K_{32} - K_{31}K_{12})(I - K_{21}K_{12})^{-1}(K_{23} - K_{21}K_{13}) \end{pmatrix} \quad (3)$$

where $\sim$ denotes matrix congruence (Bhatia, 2007), and

$$X = \begin{pmatrix} I & 0 & 0 \\ -K_{21} & I & 0 \\ (K_{32} - K_{31}K_{12})(I - K_{21}K_{12})^{-1}K_{21} - K_{31} & -(K_{32} - K_{31}K_{12})(I - K_{21}K_{12})^{-1} & I \end{pmatrix}.$$

From (3), the result of Theorem 1 for $p = 3$ reduces to $I - K_{21}K_{12}$ and $I - K_{31}K_{13} - (K_{32} - K_{31}K_{12})(I - K_{21}K_{12})^{-1}(K_{23} - K_{21}K_{13})$ being nonnegative definite. Cases for $p \geqslant 4$ can be studied similarly.

A special case of Theorem 1 particularly useful for applications is contained in the following corollary.

Corollary 1. *Let $\beta_{ij} = \beta_{ii}^{1/2} S_{ij} \beta_{jj}^{1/2}$ where the set of diagonal matrices S_{ij} $(i \neq j = 1, \ldots, p)$ is such that*

$$\begin{pmatrix} I & S_{12} & \cdots & S_{1p} \\ S_{21} & I & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ S_{p1} & \cdots & \cdots & I \end{pmatrix} \tag{4}$$

is nonnegative definite. Then β is nonnegative definite.

Defining the diagonal elements of S_{ij} as s_{ijk} $(k = 1, \ldots, n)$, the condition that (4) is nonnegative definite reduces to all matrices of the form $(s_{ijk})_{i,j=1}^{p}$ with $s_{iik} = 1$ being nonnegative definite. In particular, as with the bivariate case, $|s_{ijk}| \leqslant 1$. Hence, ensuring that the $np \times np$ matrix (4) is valid can be reduced to ensuring that a set of $p \times p$ matrices is valid.

For example, when $p = 3$, the conditions derived from (3) for diagonal matrices S_{ij} reduce to $I - S_{12}^2$ and $(I - S_{12}^2)(I - S_{13}^2) - (S_{23} - S_{12}S_{13})^2$ being nonnegative definite. Therefore, in addition to $|s_{ijk}| \leqslant 1$, one needs for example that $s_{13k} \in \{s_{12k}s_{23k} - (s_{12k}^2 s_{23k}^2 + 1 - s_{12k}^2 - s_{23k}^2)^{1/2}, s_{12k}s_{23k} + (s_{12k}^2 s_{23k}^2 + 1 - s_{12k}^2 - s_{23k}^2)^{1/2}\}$.

In practice, assuming the diagonal structure of (4) still leads to many free parameters. It will be convenient, and typically necessary, to rewrite s_{ijk} as a regression on some covariates which define how cross-correlation nonstationarity evolves across the spatial domain. For example, in the next section we use elevation to reduce the dimensionality of the parameter space for S_{12}.

If the diagonal form used in Corollary 1 is too restrictive for a given situation, then a more general construction can be derived from Theorem 1. In the following corollary, we generalize the diagonal form of (4) to a full singular value decomposition for certain classes of unitary matrices. Its proof is in the Appendix.

Corollary 2. *Given a set of arbitrary $n \times n$ unitary matrices U_i $(i = 1, \ldots, p)$, let $\beta_{ij} = \beta_{ii}^{1/2} U_i^{\mathrm{T}} S_{ij} U_j \beta_{jj}^{1/2}$ where the set of diagonal matrices S_{ij} $(i \neq j = 1, \ldots, p)$ is such that*

$$\begin{pmatrix} I & S_{12} & \cdots & S_{1p} \\ S_{21} & I & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ S_{p1} & \cdots & \cdots & I \end{pmatrix}$$

is nonnegative definite. Then β is nonnegative definite.

Corollary 2 is general in the sense that nondiagonal contraction matrices can be used.

3. Temperature fields in complex terrain

We consider a network of 145 observation stations in Colorado, making up a subset of the Global Historical Climatology Network (Peterson & Vose, 1997). Between the years 1893 and 2011, we examine minimum and maximum temperature residuals, $N(s)$ and $X(s)$ respectively, on August 1. These residuals are found by removing a station-specific mean, estimated as the arithmetic average over the 119 years.

We require an estimate of the marginal covariance functions, which we parameterize as a parsimonious bivariate Matérn covariance structure, augmented with a nugget effect (Guttorp & Gneiting, 2006; Gneiting et al., 2010), so that both minimum and maximum temperatures have distinct parameters, including distinct nugget effects. To account for terrain effects on variability, we model a local standard deviation parameter as $\sigma(s) = \exp\{\xi_0 + \xi_1 h(s)\}$, where $h(s)$ is the elevation at location s. The standard deviation parameters, Matérn smoothnesses and scale, and the nugget effect variances are estimated by least squares distance from the empirical covariance matrix, thereby imposing no distributional assumptions on the bivariate process apart from existence of first and second moments.

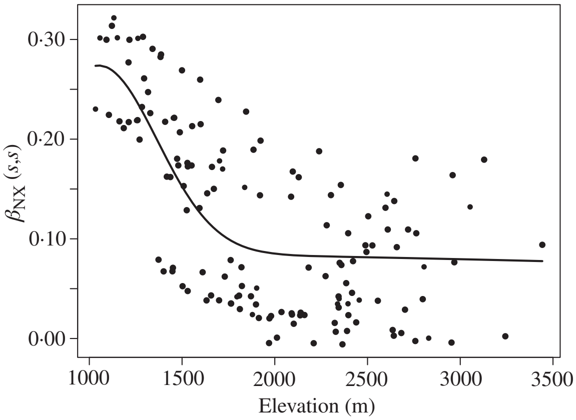

Fig. 2. Nonparametric and model co-located cross-correlation coefficients between minimum and maximum temperature residuals in Colorado on August 1 shown as dots and a curve, respectively, as a function of elevation.

The cross-correlation between minimum and maximum temperature residuals is highly dependent on terrain, and we choose to use a diagonal contraction matrix as in Proposition 1 with kth diagonal entry

$$K(s_k;\theta) = \theta_0 \exp\left[-\left\{\frac{h(s_k)-1000}{\theta_1}\right\}^2\right],$$

with 1000 m being approximately the minimal elevation in Colorado. This functional form accounts for the dependence of cross-correlation on elevation.

Write $s_1,\ldots,s_n$ for the $n=145$ observation network locations; then a nonparametric estimator of $\beta_{NX}(x,y)$ at arbitrary location pairs $x,y\in\mathbb{R}^d$, $d=2$ for our case, is

$$\hat{\beta}_{NX}(x,y) = \frac{\sum_{k=1}^n K_\lambda(\|x-s_k\|)^{1/2}K_\lambda(\|y-s_k\|)^{1/2}N(s_k)X(s_k)}{\sigma_N(x)\sigma_X(y)\{\sum_{k=1}^n K_\lambda(\|x-s_k\|)\}^{1/2}\{\sum_{k=1}^n K_\lambda(\|y-s_k\|)\}^{1/2}} \quad (5)$$

where $\sigma_N(x)$ and $\sigma_X(y)$ are the marginal standard deviation parameters for minimum and maximum temperature at locations x and y, respectively. Here, $K_\lambda(\cdot)$ is a nonnegative kernel function with bandwidth λ. This Nadaraya–Watson type estimator is available at any location pair (x,y), regardless of the observation network. The initial kernel smoothed estimate requires a bandwidth in (5); we estimate it by leave-one-out cross validation, yielding a bandwidth of $\lambda = 80$ km.

We estimate the parameters $\theta = (\theta_0,\theta_1)^{\mathrm{T}}$ via

$$\min_\theta \sum_{k,\ell=1}^n \{\beta_{NX}(s_k,s_\ell;\theta) - \hat{\beta}_{NX}(s_k,s_\ell)\}^2,$$

where $\theta_0\in[0,1]$ and $\theta_1>0$. Figure 2 shows the estimated curve of co-located cross-correlation $\beta_{NX}(s,s)$ plotted with the nonparametric estimates in (5) as a function of elevation. Our approach can simultaneously capture the positive cross-correlation at low elevations and allows minimum and maximum temperatures to be less dependent at high elevations, is available at any location pairs, and still yields a nonnegative definite covariance matrix. Figure 3 shows a map of elevations in Colorado with the corresponding spatially varying co-located cross-correlation coefficients $\beta_{NX}(s,s)$. Any covariance model can be used, and, for example, simulating from a model with a nonstationary cross-correlation coefficient will preserve the positive correlation between minimum and maximum temperature across the eastern plains of Colorado while allowing the variables to be effectively independent across the mountainous central region.

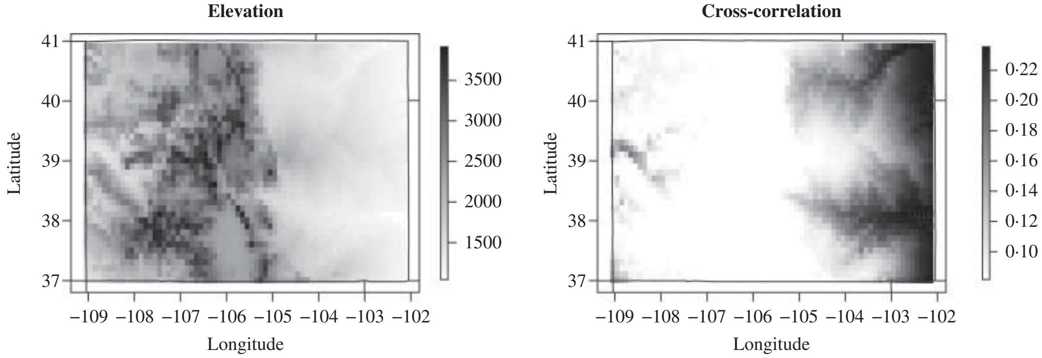

Fig. 3. Map of elevations in Colorado in metres with estimated co-located cross-correlation between minimum and maximum temperature residuals for August 1.

Acknowledgement

This research was partially supported by an award made by the King Abdullah University of Science and Technology and by the National Science Foundation.

Appendix

Proof of Proposition 1. The proof closely follows Proposition 1.3.2 of Bhatia (2007). We use $\sim$ to denote matrix congruence, and let $\beta_{11}^{-1/2}\beta_{12}\beta_{22}^{-1/2} = USV$ be its singular value decomposition. Then

$$\begin{pmatrix}\beta_{11} & \beta_{12}\\ \beta_{21} & \beta_{22}\end{pmatrix} \sim \begin{pmatrix}\beta_{11}^{-1/2} & 0\\ 0 & \beta_{22}^{-1/2}\end{pmatrix}\begin{pmatrix}\beta_{11} & \beta_{12}\\ \beta_{21} & \beta_{22}\end{pmatrix}\begin{pmatrix}\beta_{11}^{-1/2} & 0\\ 0 & \beta_{22}^{-1/2}\end{pmatrix} = \begin{pmatrix} I & \beta_{11}^{-1/2}\beta_{12}\beta_{22}^{-1/2}\\ \beta_{22}^{-1/2}\beta_{12}^{\mathrm{T}}\beta_{11}^{-1/2} & I\end{pmatrix}$$

$$= \begin{pmatrix} I & USV\\ V^{\mathrm{T}}SU^{\mathrm{T}} & I\end{pmatrix} = \begin{pmatrix} U & 0\\ 0 & V^{\mathrm{T}}\end{pmatrix}\begin{pmatrix} I & S\\ S & I\end{pmatrix}\begin{pmatrix} U^{\mathrm{T}} & 0\\ 0 & V\end{pmatrix}$$

where 0 is a matrix of zeroes and I is the identity matrix, both of dimension $n \times n$. By matrix congruence, β_{12} is nonnegative definite if and only if

$$\begin{pmatrix} I & S\\ S & I\end{pmatrix}$$

is nonnegative definite, that is, all elements of S are bounded by one. In particular, $K = \beta_{11}^{-1/2}\beta_{12}\beta_{22}^{-1/2}$ is a contraction matrix. Rearranging terms yields the result. □

Proof of Theorem 1. The proof is similar to that of Proposition 1. We have

$$\begin{pmatrix}\beta_{11} & \cdots & \beta_{1p}\\ \vdots & \ddots & \vdots\\ \beta_{p1} & \cdots & \beta_{pp}\end{pmatrix} \sim \begin{pmatrix}\beta_{11}^{-1/2} & 0 & \cdots & 0\\ 0 & \beta_{22}^{-1/2} & \ddots & \vdots\\ \vdots & \ddots & \ddots & \vdots\\ 0 & \cdots & \cdots & \beta_{pp}^{-1/2}\end{pmatrix}\begin{pmatrix}\beta_{11} & \cdots & \beta_{1p}\\ \vdots & \ddots & \vdots\\ \beta_{p1} & \cdots & \beta_{pp}\end{pmatrix}\begin{pmatrix}\beta_{11}^{-1/2} & 0 & \cdots & 0\\ 0 & \beta_{22}^{-1/2} & \ddots & \vdots\\ \vdots & \ddots & \ddots & \vdots\\ 0 & \cdots & \cdots & \beta_{pp}^{-1/2}\end{pmatrix}$$

$$= \begin{pmatrix} I & \beta_{11}^{-1/2}\beta_{12}\beta_{22}^{-1/2} & \cdots & \beta_{11}^{-1/2}\beta_{1p}\beta_{pp}^{-1/2}\\ \beta_{22}^{-1/2}\beta_{21}\beta_{11}^{-1/2} & I & \ddots & \vdots\\ \vdots & \ddots & \ddots & \vdots\\ \beta_{pp}^{-1/2}\beta_{p1}\beta_{11}^{-1/2} & \cdots & \cdots & I\end{pmatrix},$$

where I is the $n \times n$ identity matrix and 0 is a $n \times n$ matrix of zeroes. By matrix congruence, β is nonnegative definite if and only if K is nonnegative definite. □

Proof of Corollary 2. From the proof of Theorem 1, it suffices to show that the following matrix is nonnegative definite,

$$\begin{pmatrix} I & \beta_{11}^{-1/2}\beta_{12}\beta_{22}^{-1/2} & \cdots & \beta_{11}^{-1/2}\beta_{1p}\beta_{pp}^{-1/2} \\ \beta_{22}^{-1/2}\beta_{21}\beta_{11}^{-1/2} & I & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \beta_{pp}^{-1/2}\beta_{p1}\beta_{11}^{-1/2} & \cdots & \cdots & I \end{pmatrix} = \begin{pmatrix} I & U_1^{\mathrm{T}}S_{12}U_2 & \cdots & U_1^{\mathrm{T}}S_{1p}U_p \\ U_2^{\mathrm{T}}S_{21}U_1 & I & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ U_p^{\mathrm{T}}S_{p1}U_1 & \cdots & \cdots & I \end{pmatrix}$$

$$= \begin{pmatrix} U_1^{\mathrm{T}} & 0 & \cdots & 0 \\ 0 & U_2^{\mathrm{T}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & U_p^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} I & S_{12} & \cdots & S_{1p} \\ S_{21} & I & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ S_{p1} & \cdots & \cdots & I \end{pmatrix} \begin{pmatrix} U_1 & 0 & \cdots & 0 \\ 0 & U_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & U_p \end{pmatrix}.$$

By matrix congruence, the result is proven. □

REFERENCES

APANASOVICH, T. V. & GENTON, M. G. (2010). Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika* **97**, 15–30.

APANASOVICH, T. V., GENTON, M. G. & SUN, Y. (2012). A valid Matérn class of cross-covariance functions for multivariate random fields with any number of components. *J. Am. Statist. Assoc.* **107**, 180–93.

BHATIA, R. (2007). *Positive Definite Matrices*. Princeton: Princeton Press.

FANSHAWE, T. R. & DIGGLE, P. J. (2012). Bivariate geostatistical modelling: A review and an application to spatial variation in radon concentrations. *Envir. Ecol. Statist.* **19**, 139–60.

FURRER, R. & GENTON, M. G. (2011). Aggregation-cokriging for highly-multivariate spatial data. *Biometrika* **98**, 615–31.

GASPARI, G. & COHN, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quart. J. R. Meteorol. Soc.* **125**, 723–57.

GELFAND, A. E., SCHMIDT, A. M., BANERJEE, S. & SIRMANS, C. F. (2004). Nonstationary multivariate process modeling through spatially varying coregionalization (with discussion and rejoinder). *Test* **13**, 263–312.

GNEITING, T., KLEIBER, W. & SCHLATHER, M. (2010). Matérn cross-covariance functions for multivariate random fields. *J. Am. Statist. Assoc.* **105**, 1167–77.

GOULARD, M. & VOLTZ, M. (1992). Linear coregionalization model: Tools for estimation and choice of cross-variogram matrix. *Math. Geol.* **24**, 269–82.

GUTTORP, P. & GNEITING, T. (2006). Studies in the history of probability and statistics XLIX: On the Matérn correlation family. *Biometrika* **93**, 989–95.

KLEIBER, W. & NYCHKA, D. (2012). Nonstationary modeling for multivariate spatial processes. *J. Mult. Anal.* **112**, 76–91.

LI, B. & ZHANG, H. (2011). An approach to modeling asymmetric multivariate spatial covariance structures. *J. Mult. Anal.* **102**, 1445–53.

MAJUMDAR, A. & GELFAND, A. E. (2007). Multivariate spatial modeling for geostatistical data using convolved covariance functions. *Math. Geol.* **39**, 225–45.

MAJUMDAR, A., PAUL, D. & BAUTISTA, D. (2010). A generalized convolution model for multivariate nonstationary spatial processes. *Statist. Sinica* **20**, 675–95.

PETERSON, T. C. & VOSE, R. S. (1997). An overview of the Global Historical Climatology Network temperature database. *Bull. Am. Meteor. Soc.* **78**, 2837–49.

PORCU, E. & ZASTAVNYI, V. (2011). Characterization theorems for some classes of covariance functions associated to vector valued random fields. *J. Mult. Anal.* **102**, 1293–301.

SCHMIDT, A. M. & GELFAND, A. E. (2003). A Bayesian coregionalization approach for multivariate pollutant data. *J. Geophys. Res. – Atmos.* **108**, 8783–90.

VER HOEF, J. M. & BARRY, R. P. (1998). Constructing and fitting models for cokriging and multivariable spatial prediction. *J. Statist. Plan. Infer.* **69**, 275–94.

WACKERNAGEL, H. (2003). *Multivariate Geostatistics*. Berlin: Springer-Verlag, 3rd ed.

ZHANG, H. (2007). Maximum-likelihood estimation for multivariate spatial linear coregionalization models. *Environmetrics* **18**, 125–39.

[*Received May* 2012. *Revised August* 2012]