Cross-Covariance Functions for Multivariate Geostatistics[1]

Marc G. Genton and William Kleiber

Abstract. Continuously indexed datasets with multiple variables have become ubiquitous in the geophysical, ecological, environmental and climate sciences, and pose substantial analysis challenges to scientists and statisticians. For many years, scientists developed models that aimed at capturing the spatial behavior for an individual process; only within the last few decades has it become commonplace to model multiple processes jointly. The key difficulty is in specifying the cross-covariance function, that is, the function responsible for the relationship between distinct variables. Indeed, these cross-covariance functions must be chosen to be consistent with marginal covariance functions in such a way that the second-order structure always yields a nonnegative definite covariance matrix. We review the main approaches to building cross-covariance models, including the linear model of coregionalization, convolution methods, the multivariate Matérn and nonstationary and space–time extensions of these among others. We additionally cover specialized constructions, including those designed for asymmetry, compact support and spherical domains, with a review of physics-constrained models. We illustrate select models on a bivariate regional climate model output example for temperature and pressure, along with a bivariate minimum and maximum temperature observational dataset; we compare models by likelihood value as well as via cross-validation co-kriging studies. The article closes with a discussion of unsolved problems.

Key words and phrases: Asymmetry, co-kriging, multivariate random fields, nonstationarity, separability, smoothness, spatial statistics, symmetry.

1. INTRODUCTION

1.1 Motivation

The occurrence of multivariate data indexed by spatial coordinates in a large number of applications has prompted sustained interest in statistics in recent years. For instance, in environmental and climate sciences, monitors collect information on multiple variables such as temperature, pressure, wind speed and direction and various pollutants. Similarly, the output of climate models generate multiple variables, and there are multiple distinct climate models. Physical models in computer experiments often involve multiple processes that are indexed by not only space and time, but also parameter settings. With the increasing availability and scientific interest in multivariate processes, statistical science faces new challenges and an expanding horizon of opportunities for future exploration.

Geostatistical applications mainly focus on interpolation, simulation or statistical modeling. Interpolation or smoothing in spatial statistics usually is synonymous with kriging, the best linear unbiased prediction under squared loss (Cressie, 1993). With multiple variables, interpolation becomes a multivariate problem, and is traditionally accommodated via co-kriging, the multi-

Marc G. Genton is Professor, CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia (e-mail: marc.genton@kaust.edu.sa). William Kleiber is Assistant Professor, Department of Applied Mathematics, University of Colorado, Boulder, Colorado 80309-0526, USA (e-mail: william.kleiber@colorado.edu).

variate extension of kriging. Co-kriging is often particularly useful when one variable is of primary importance, but is correlated with other types of processes that are more readily observed (Almeida and Journel, 1994; Wackernagel, 1994; Journel, 1999; Shmaryan and Journel, 1999; Subramanyam and Pandalai, 2008). Much expository work has been developed on co-kriging, see Myers (1982, 1983, 1991, 1992), Long and Myers (1997), Furrer and Genton (2011) and Sang, Jun and Huang (2011) for discussion and technical details.

Consider a p-dimensional multivariate random field $\mathbf{Z}(\mathbf{s}) = \{Z_1(\mathbf{s}), \ldots, Z_p(\mathbf{s})\}^{\mathrm{T}}$ defined on $\mathbb{R}^d$, $d \geq 1$, where $Z_i(\mathbf{s})$ is the ith process at location $\mathbf{s}$, for $i = 1, \ldots, p$. If $\mathbf{Z}(\mathbf{s})$ is assumed to be a Gaussian multivariate random field, then only its mean vector $\boldsymbol{\mu}(\mathbf{s}) = \mathrm{E}\{\mathbf{Z}(\mathbf{s})\}$ and cross-covariance matrix function $\mathbf{C}(\mathbf{s}_1, \mathbf{s}_2) = \mathrm{cov}\{\mathbf{Z}(\mathbf{s}_1), \mathbf{Z}(\mathbf{s}_2)\} = \{C_{ij}(\mathbf{s}_1, \mathbf{s}_2)\}_{i,j=1}^p$ composed of functions

$$(1)\quad C_{ij}(\mathbf{s}_1, \mathbf{s}_2) = \mathrm{cov}\{Z_i(\mathbf{s}_1), Z_j(\mathbf{s}_2)\}, \quad \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d,$$

for $i, j = 1, \ldots, p$, need to be described to fully specify the multivariate random field. Authors typically refer to C_{ij} as direct- or marginal-covariance functions for $i = j$, and cross-covariance functions for $i \neq j$. Here, we assume that $\mathbf{Z}(\mathbf{s})$ is a mean zero process. The quantities $\rho_{ij}(\mathbf{s}_1, \mathbf{s}_2) = C_{ij}(\mathbf{s}_1, \mathbf{s}_2)/\{C_{ii}(\mathbf{s}_1, \mathbf{s}_1) C_{jj}(\mathbf{s}_2, \mathbf{s}_2)\}^{1/2}$ are the cross-correlation functions. Our goal is then to construct valid and flexible cross-covariance functions (1), that is, the matrix-valued mapping $\mathbf{C}: \mathbb{R}^d \times \mathbb{R}^d \to M_{p \times p}$, where $M_{p \times p}$ is the set of $p \times p$ real-valued matrices, must be nonnegative definite in the following sense. The covariance matrix $\boldsymbol{\Sigma}$ of the random vector $\{\mathbf{Z}(\mathbf{s}_1)^{\mathrm{T}}, \ldots, \mathbf{Z}(\mathbf{s}_n)^{\mathrm{T}}\}^{\mathrm{T}} \in \mathbb{R}^{np}$:

$$(2)\quad \boldsymbol{\Sigma} = \begin{pmatrix} \mathbf{C}(\mathbf{s}_1, \mathbf{s}_1) & \mathbf{C}(\mathbf{s}_1, \mathbf{s}_2) & \cdots & \mathbf{C}(\mathbf{s}_1, \mathbf{s}_n) \\ \mathbf{C}(\mathbf{s}_2, \mathbf{s}_1) & \mathbf{C}(\mathbf{s}_2, \mathbf{s}_2) & \cdots & \mathbf{C}(\mathbf{s}_2, \mathbf{s}_n) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}(\mathbf{s}_n, \mathbf{s}_1) & \mathbf{C}(\mathbf{s}_n, \mathbf{s}_2) & \cdots & \mathbf{C}(\mathbf{s}_n, \mathbf{s}_n) \end{pmatrix},$$

should be nonnegative definite: $\mathbf{a}^{\mathrm{T}} \boldsymbol{\Sigma} \mathbf{a} \geq 0$ for any vector $\mathbf{a} \in \mathbb{R}^{np}$, any spatial locations $\mathbf{s}_1, \ldots, \mathbf{s}_n$, and any integer n. Fanshawe and Diggle (2012) reviewed approaches for the bivariate case $p = 2$, although most techniques can be readily extended to $p > 2$, and Álvarez, Rosasco and Lawrence (2012) reviewed approaches for machine learning.

A multivariate random field is second-order stationary (or just stationary) if the marginal and cross-covariance functions depend only on the separation vector $\mathbf{h} = \mathbf{s}_1 - \mathbf{s}_2$, that is, there is a mapping $C_{ij}: \mathbb{R}^d \to \mathbb{R}$ such that

$$\mathrm{cov}\{Z_i(\mathbf{s}_1), Z_j(\mathbf{s}_2)\} = C_{ij}(\mathbf{h}), \quad \mathbf{h} \in \mathbb{R}^d.$$

Otherwise, the process is nonstationary. Stationarity can be thought of as an invariance property under the translation of coordinates. A test for the stationarity of a multivariate random field can be found in Jun and Genton (2012).

A multivariate random field is isotropic if it is stationary and invariant under rotations and reflections, that is, there is a mapping $C_{ij}: \mathbb{R}_+ \cup \{0\} \to \mathbb{R}$ such that

$$\mathrm{cov}\{Z_i(\mathbf{s}_1), Z_j(\mathbf{s}_2)\} = C_{ij}(\|\mathbf{h}\|), \quad \mathbf{h} \in \mathbb{R}^d,$$

where $\| \cdot \|$ denotes the Euclidean norm. Otherwise, the multivariate random field is anisotropic. Isotropy or even stationarity are not always realistic, especially for large spatial regions, but sometimes are satisfactory working assumptions and serve as basic elements of more sophisticated anisotropic and nonstationary models.

In the univariate setting, variograms are often the main focus in geostatistics, and are defined as the variance of contrasts. Variograms can be extended to multivariate random fields in two ways: A covariance-based cross-variogram (Myers, 1982) defined as

$$(3)\quad \mathrm{cov}\{Z_i(\mathbf{s}_1) - Z_i(\mathbf{s}_2), Z_j(\mathbf{s}_1) - Z_j(\mathbf{s}_2)\},$$

$\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d$, and a variance-based cross-variogram (Myers, 1991), also coined pseudo cross-variogram,

$$(4)\quad \mathrm{var}\{Z_i(\mathbf{s}_1) - Z_j(\mathbf{s}_2)\}, \quad \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d.$$

The corresponding stationary versions are immediate. Cressie and Wikle (1998) reviewed the differences between (3) and (4), and argued that (4) is more appropriate for co-kriging because it yields the same optimal co-kriging predictor as the one obtained with the cross-covariance function C_{ij} in (1); see also Ver Hoef and Cressie (1993) and Huang, Yao, Cressie and Hsing (2009). Unfortunately, the interpretation of cross-variograms is difficult, and so most authors favor working with covariance and cross-covariance formulations.

1.2 Properties of Cross-Covariance Matrix Functions

Because the covariance matrix $\boldsymbol{\Sigma}$ in (2) must be symmetric, the matrix functions must satisfy $\mathbf{C}(\mathbf{s}_1, \mathbf{s}_2) = \mathbf{C}(\mathbf{s}_2, \mathbf{s}_1)^{\mathrm{T}}$, or $\mathbf{C}(\mathbf{h}) = \mathbf{C}(-\mathbf{h})^{\mathrm{T}}$ under stationarity. Therefore, cross-covariance matrix functions are not symmetric in general, that is,

$$C_{ij}(\mathbf{s}_1, \mathbf{s}_2) = \mathrm{cov}\{Z_i(\mathbf{s}_1), Z_j(\mathbf{s}_2)\}$$
$$\neq \mathrm{cov}\{Z_j(\mathbf{s}_1), Z_i(\mathbf{s}_2)\} = C_{ji}(\mathbf{s}_1, \mathbf{s}_2),$$

$\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d$, unless the cross-covariance functions themselves are all symmetric (Wackernagel, 2003).

However, the collocated matrices $\mathbf{C}(\mathbf{s}, \mathbf{s})$, or $\mathbf{C}(\mathbf{0})$ under stationarity, are symmetric and nonnegative definite.

The marginal and cross-covariance functions satisfy $|C_{ij}(\mathbf{s}_1, \mathbf{s}_2)|^2 \leq C_{ii}(\mathbf{s}_1, \mathbf{s}_1) C_{jj}(\mathbf{s}_2, \mathbf{s}_2)$, or $|C_{ij}(\mathbf{h})|^2 \leq C_{ii}(\mathbf{0}) C_{jj}(\mathbf{0})$ under stationarity. However, $|C_{ij}(\mathbf{s}_1, \mathbf{s}_2)|$ need not be less than or equal to $C_{ij}(\mathbf{s}_1, \mathbf{s}_1)$, or $|C_{ij}(\mathbf{h})|$ need not be less than or equal to $C_{ij}(\mathbf{0})$ under stationarity. This is because the maximum value of $C_{ij}(\mathbf{h})$ is not restricted to occur at $\mathbf{h} = \mathbf{0}$, unless $i = j$, and in fact this sometimes occurs in practice (Li and Zhang, 2011). Thus, there are no similar bounds between $|C_{ij}(\mathbf{s}_1, \mathbf{s}_2)|^2$ and $C_{ii}(\mathbf{s}_1, \mathbf{s}_2) C_{jj}(\mathbf{s}_1, \mathbf{s}_2)$, or between $|C_{ij}(\mathbf{h})|^2$ and $C_{ii}(\mathbf{h}) C_{jj}(\mathbf{h})$ under stationarity.

A cross-covariance matrix function is separable if

$$(5) \quad C_{ij}(\mathbf{s}_1, \mathbf{s}_2) = \rho(\mathbf{s}_1, \mathbf{s}_2) R_{ij}, \quad \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d,$$

for all $i, j = 1, \ldots, p$, where $\rho(\mathbf{s}_1, \mathbf{s}_2)$ is a valid, nonstationary or stationary, correlation function and $R_{ij} = \mathrm{cov}(Z_i, Z_j)$ is the nonspatial covariance between variables i and j. Mardia and Goodall (1993) introduced and used separability to model multivariate spatio-temporal data, and Bhat, Haran and Goes (2010) used separable covariances in the context of computer model calibration. In the past, separable cross-covariance structures were sometimes called intrinsic coregionalizations (Helterbrand and Cressie, 1994).

With a large number of processes, detecting structures of the multivariate random process such as symmetry and separability can be difficult via elementary data analytic techniques. Li, Genton and Sherman (2008) proposed an approach based on the asymptotic distribution of the sample cross-covariance estimator to test these various structures. Their methodology allows the practitioner to assess the underlying dependence structure of the data and to suggest appropriate cross-covariance functions, an important part of model building.

In the special case of stationary matrix-valued covariance functions, there is an intimate link between the cross-covariance matrix function and its spectral representation. In particular, define the cross-spectral densities $f_{ij}: \mathbb{R}^d \to \mathbb{R}$ as

$$f_{ij}(\boldsymbol{\omega}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-\iota \mathbf{h}^{\mathrm{T}} \boldsymbol{\omega}} C_{ij}(\mathbf{h}) \, d\mathbf{h}, \quad \boldsymbol{\omega} \in \mathbb{R}^d,$$

where $\iota = \sqrt{-1}$ is the imaginary number. A necessary and sufficient condition for $\mathbf{C}(\cdot)$ to be a valid (i.e., nonnegative definite), stationary matrix-valued covariance function is for the matrix function $\mathbf{f}(\boldsymbol{\omega}_0) = \{f_{ij}(\boldsymbol{\omega}_0)\}_{i,j=1}^p$ to be nonnegative definite for any $\boldsymbol{\omega}_0$

(Cramér, 1940). While Cramér's original result is stated in terms of measures of bounded variation, in practice using spectral densities is preferred. This can be viewed as a multivariate extension of Bochner's celebrated theorem (Bochner, 1955). The analogue of Schoenberg's theorem for multivariate random fields, that is, Bochner's theorem for isotropic cross-covariance functions, has recently been investigated by Alonso-Malaver, Porcu and Giraldo (2013, 2015).

1.3 Estimation of Cross-Covariances

The empirical estimator of the cross-covariance matrix function of a stationary multivariate random field is

$$
\widehat{\mathbf{C}}(\mathbf{h}) = \frac{1}{|N(\mathbf{h})|} \sum_{(k,l) \in N(\mathbf{h})} \{\mathbf{Z}(\mathbf{s}_k) - \bar{\mathbf{Z}}\}
$$

(6)

$$
\cdot \{\mathbf{Z}(\mathbf{s}_l) - \bar{\mathbf{Z}}\}^{\mathrm{T}},
$$

$\mathbf{h} \in \mathbb{R}^d$, where $N(\mathbf{h}) = \{(k, l) | \mathbf{s}_k - \mathbf{s}_l = \mathbf{h}\}$, $|N(\mathbf{h})|$ denotes its cardinality, and $\bar{\mathbf{Z}} = \frac{1}{n} \sum_{k=1}^n \mathbf{Z}(\mathbf{s}_k)$ is the sample mean vector. A valid parametric model is then typically fit by least squares methods to the empirical estimates in (6). Alternatively, one can use likelihood-based methods or the Bayesian paradigm (Brown, Le and Zidek, 1994). In any case, valid and flexible cross-covariance models are needed. Künsch, Papritz and Bassi (1997) studied generalized cross-covariances and their estimation.

Papritz, Künsch and Webster (1993) discussed empirical estimators of the cross-variogram (3) and (4). Unlike the pseudo cross-variogram, the cross-variogram (3) has the disadvantage that it cannot be estimated when the variables are not observed at the same spatial locations. Lark (2003) proposed two outlier-robust estimators of the pseudo cross-variogram (4) and applied them in a multivariate geostatistical analysis of soil properties. Furrer (2005) studied the bias of the empirical cross-covariance matrix $\mathbf{C}(\mathbf{0})$ estimation under spatial dependence using both fixed-domain and increasing-domain asymptotics. Lim and Stein (2008) investigated a spectral approach based on spatial cross-periodograms for data on a lattice and studied their properties using fixed-domain asymptotics.

2. CROSS-COVARIANCES BUILT FROM UNIVARIATE MODELS

The most common approach to building cross-covariance functions is by combining univariate covariance functions. The three main options in this vein are the linear model of coregionalization, various convolution techniques and the use of latent dimensions.

2.1 Linear Model of Coregionalization

Probably the most popular approach of combining univariate covariances is the so-called linear model of coregionalization (LMC) for stationary random fields (Bourgault and Marcotte, 1991; Goulard and Voltz, 1992; Grzebyk and Wackernagel, 1994; Vargas-Guzmán, Warrick and Myers, 2002; Schmidt and Gelfand, 2003; Wackernagel, 2003). It consists of representing the multivariate random field as a linear combination of r independent univariate random fields. The resulting cross-covariance functions take the form

$$(7) \qquad C_{ij}(\mathbf{h}) = \sum_{k=1}^{r} \rho_k(\mathbf{h}) A_{ik} A_{jk}, \quad \mathbf{h} \in \mathbb{R}^d,$$

for an integer $1 \le r \le p$, where $\rho_k(\cdot)$ are valid stationary correlation functions and $\mathbf{A} = (A_{ij})_{i,j=1}^{p,r}$ is a $p \times r$ full rank matrix. When $r = 1$, the cross-covariance function (7) is separable as in (5). The allure of this approach is that only r univariate covariances $\rho_k(\mathbf{h})$ must be specified, thus avoiding direct specification of a valid cross-covariance matrix function. The LMC can additionally be built from a conditional perspective (Royle and Berliner, 1999; Gelfand et al., 2004). Note that the discrete sum representation (7) can also be interpreted as a scale mixture (Porcu and Zastavnyi, 2011).

With a large number of processes, the number of parameters can quickly become unwieldy and the resulting estimation difficult. Zhang (2007) described maximum likelihood estimation of the spatial LMC based on an EM algorithm, whereas Schmidt and Gelfand (2003) proposed a Bayesian coregionalization approach with application to multivariate pollutant data. A second drawback of the LMC is that the smoothness of any component of the multivariate random field is restricted to that of the roughest underlying univariate process.

2.2 Convolution Methods

Convolution methods fall into the two categories of kernel and covariance convolution. The kernel convolution method (Ver Hoef and Barry, 1998; Ver Hoef, Cressie and Barry, 2004) uses

$$C_{ij}(\mathbf{h})$$
$$= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} k_i(\mathbf{v}_1) k_j(\mathbf{v}_2) \rho(\mathbf{v}_1 - \mathbf{v}_2 + \mathbf{h}) \, d\mathbf{v}_1 \, d\mathbf{v}_2,$$

$\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d$, where the k_i are square integrable kernel functions and $\rho(\cdot)$ is a valid stationary correlation function. This approach assumes that all the spatial processes $Z_i(\mathbf{s})$, for $i = 1, \ldots, p$, are generated by the same underlying process, which is very restrictive in that it imposes strong dependence between all constituent processes $Z_i(\mathbf{s})$. Overall, this approach and its parameters can be difficult to interpret and, except for some special cases, requires numerical integration.

Covariance convolution for stationary spatial random fields (Gaspari and Cohn, 1999; Gaspari et al., 2006; Majumdar and Gelfand, 2007) yields

$$C_{ij}(\mathbf{h}) = \int_{\mathbb{R}^d} C_i(\mathbf{h} - \mathbf{k}) C_j(\mathbf{k}) \, d\mathbf{k}, \quad \mathbf{h} \in \mathbb{R}^d,$$

where C_i are square integrable functions. Although some closed-form expressions exist, this method usually requires numerical integration. A particularly useful example of a closed form solution is when the C_i are Matérn correlation functions with common scale parameters. In this setup, Matérn correlations are closed under convolution and this approach results in a special case of the multivariate Matérn model (Gneiting, Kleiber and Schlather, 2010).

2.3 Latent Dimensions

Another approach to build valid cross-covariance functions based on univariate ($p = 1$) spatial covariances was put forward by Apanasovich and Genton (2010) (see also Porcu and Zastavnyi, 2011). Their idea was to create additional latent dimensions that represent the various variables to be modeled. Specifically, each component i of the multivariate random field $\mathbf{Z}(\mathbf{s})$ is represented as a point $\boldsymbol{\xi}_i = (\xi_{i1}, \ldots, \xi_{ik})^{\mathrm{T}}$ in $\mathbb{R}^k$, $i = 1, \ldots, p$, for an integer $1 \le k \le p$, yielding the marginal and cross-covariance functions

$$(8) \quad C_{ij}(\mathbf{s}_1, \mathbf{s}_2) = C\{(\mathbf{s}_1, \boldsymbol{\xi}_i), (\mathbf{s}_2, \boldsymbol{\xi}_j)\}, \quad \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d,$$

where C is a valid univariate covariance function on $\mathbb{R}^{d+k}$; see Gneiting, Genton and Guttorp (2007) for a review of possible univariate covariance functions. It is immediate that the resulting cross-covariance matrix $\boldsymbol{\Sigma}$ in (2) is nonnegative definite because its entries are defined through a valid univariate covariance. If the covariance C is from a stationary or isotropic univariate random field, then so is also the cross-covariance function (8); for instance, $C_{ij}(\mathbf{h}) = C(\mathbf{h}, \boldsymbol{\xi}_i - \boldsymbol{\xi}_j)$.

As an example of the aforementioned construction, Apanasovich and Genton (2010) suggested

$$
\begin{aligned}
C_{ij}(\mathbf{h}) &= \frac{\sigma_i \sigma_j}{\|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\| + 1} \exp\left\{ \frac{-\alpha \|\mathbf{h}\|}{(\|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\| + 1)^{\beta/2}} \right\} \\
&\quad + \tau^2 I(i = j) I(\mathbf{h} = \mathbf{0}), \quad \mathbf{h} \in \mathbb{R}^d,
\end{aligned}
$$
(9)

where $I(\cdot)$ is the indicator function, $\sigma_i > 0$ are marginal standard deviations, $\tau \geq 0$ is a nugget effect, and $\alpha > 0$ is a length scale. Here, $\beta \in [0, 1]$ controls the nonseparability between space and variables, with $\beta = 0$ being the separable case. The parameters of the model are estimated by maximum likelihood or composite likelihood methods. Apanasovich and Genton (2010) provided an application to a trivariate pollution dataset from California. Further use of latent dimensions for multivariate spatio-temporal random fields are discussed in Section 7.2. The idea of latent dimensions was recently extended to modeling nonstationary processes by Bornn, Shaddick and Zidek (2012).

3. MATÉRN CROSS-COVARIANCE FUNCTIONS

The Matérn class of positive definite functions has become the standard covariance model for univariate fields (Gneiting and Guttorp, 2006). The popularity in large part is due to the work of Stein (1999) who showed that the behavior of the covariance function near the origin has fundamental implications on predictive distributions, particularly predictive uncertainty. The key feature of the Matérn is the inclusion of a smoothness parameter that directly controls correlation at small distances. The Matérn correlation function is

$$M(\mathbf{h}|\nu, a) = \frac{2^{1-\nu}}{\Gamma(\nu)}(a\|\mathbf{h}\|)^{\nu}K_{\nu}(a\|\mathbf{h}\|), \quad \mathbf{h} \in \mathbb{R}^d,$$

where K_{ν} is a modified Bessel function of order ν, $a > 0$ is a length scale parameter that controls the rate of decay of correlation at larger distances, while $\nu > 0$ is the smoothness parameter that controls behavior of correlation near the origin. The smoothness parameter is aptly named as it implies levels of mean square differentiability of the random process, with large ν yielding very smooth processes that are many times differentiable, and small ν yielding rough processes; in fact there is a direct connection between the smoothness parameter and the Hausdorff dimension of the resulting random process (Goff and Jordan, 1988).

Due to its popularity for univariate modeling, there is interest in being able to simultaneously model multiple processes, each of which marginally has a Matérn correlation structure. To this end, Gneiting, Kleiber and Schlather (2010) introduced the so-called multivariate Matérn model, where each constituent process is allowed a marginal Matérn correlation, with Matérns also composing the cross-correlation structures. In particular, the multivariate Matérn implies

$$(10) \quad \begin{aligned} \rho_{ii}(\mathbf{h}) &= M(\mathbf{h}|\nu_i, a_i) \quad \text{and} \\ \rho_{ij}(\mathbf{h}) &= \beta_{ij}M(\mathbf{h}|\nu_{ij}, a_{ij}), \quad \mathbf{h} \in \mathbb{R}^d. \end{aligned}$$

Of course, this correlation structure can be coerced to a covariance structure by multiplying $C_{ii}(\mathbf{h})$ by σ_i^2 and $C_{ij}(\mathbf{h})$ by $\sigma_i\sigma_j$. Here, β_{ij} is a collocated cross-correlation coefficient, and represents the strength of correlation between Z_i and Z_j at the same location, $\mathbf{h} = \mathbf{0}$.

The difficulty in (10) is deriving conditions on model parameters $\nu_i, \nu_{ij}, a_i, a_{ij}$ and β_{ij} that result in a valid, that is, a nonnegative definite multivariate covariance class. In the original work, Gneiting, Kleiber and Schlather (2010) described two main models, the *parsimonious Matérn* and the *full bivariate Matérn*. The parsimonious Matérn is a reduction in complexity over (10) in that $a_i = a_{ij} = a$ are held at the same value for all marginal and cross-covariances, and the cross-smoothnesses are set to the arithmetic average of the marginals, $\nu_{ij} = (\nu_i + \nu_j)/2$. The model is then valid with an easy-to-check condition on the cross-correlation coefficient β_{ij}.

The flexibility of the parsimonious Matérn is in allowing each process to have a distinct marginal smoothness behavior, and thus allowing for simultaneous modeling of highly smooth and rough fields. The natural extension to allow distinct process-dependent length scale parameters a_i turns out to be more involved. The full bivariate Matérn of Gneiting, Kleiber and Schlather (2010) allows for distinct smoothness and scale parameters for two processes (and in fact results in a characterization for $p = 2$). A second set of authors, Apanasovich, Genton and Sun (2012), were able to overcome the deficiencies of the parsimonious formulation for $p > 2$, introducing the *flexible Matérn*. The flexible Matérn works for any number of processes, allowing for each process to have distinct smoothness and scale parameters, and is as close in spirit to allowing entirely free marginal Matérn covariances with some level of cross-process dependence as is currently available. A number of simpler sufficient conditions are available by using scale mixtures (Reisert and Burkhardt, 2007; Gneiting, Kleiber and Schlather, 2010; Schlather, 2010; Porcu and Zastavnyi, 2011).

It is worth pointing out that the experimental results of both sets of authors, Gneiting, Kleiber and Schlather (2010) and Apanasovich, Genton and Sun (2012), highlighted the importance of allowing for highly flexible and distinct marginal covariance structures, while still allowing for some degree of cross-process correlation, and indeed the improvement over an independence assumption was substantial.

4. NONSTATIONARY CROSS-COVARIANCE FUNCTIONS

Geophysical, environmental and ecological spatial processes often exhibit spatial dependence that depends on fixed geographical features such as terrain or land use type, or dynamical environments such as prevailing winds. In either case, the evolving nature of spatial dependence is not well captured by stationary models, and thus the availability of nonstationary constructions is desired, that is, models such that the marginal and cross-covariance functions are now dependent on the spatial location pair, not just the lag vector, $\mathrm{cov}\{Z_i(\mathbf{s}_1), Z_j(\mathbf{s}_2)\} = C_{ij}(\mathbf{s}_1, \mathbf{s}_2)$.

Many of the aforementioned models have been extended to the nonstationary setup, including the original stationary models as special cases. The first natural extension to allowing the LMC to be nonstationary is to let the latent univariate correlations be nonstationary, so that

$$C_{ij}(\mathbf{s}_1, \mathbf{s}_2) = \sum_{k=1}^{r} \rho_k(\mathbf{s}_1, \mathbf{s}_2) A_{ik} A_{jk}, \quad \mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d,$$

where now ρ_k are nonstationary univariate correlation functions. The onus of deriving a matrix-valued nonstationary covariance function is then alleviated in favor of opting for univariate nonstationary correlations, of which there are many choices (e.g., Sampson and Guttorp, 1992; Fuentes, 2002; Paciorek and Schervish, 2006; Bornn, Shaddick and Zidek, 2012). Although this extension seems straightforward, we are unaware of any authors who have implemented such an approach. The second way to extend the LMC to a nonstationary setup is to allow the coefficients to be spatially varying (Gelfand et al., 2004), so that

$$C_{ij}(\mathbf{s}_1, \mathbf{s}_2) = \sum_{k=1}^{r} \rho_k(\mathbf{s}_1 - \mathbf{s}_2) A_{ik}(\mathbf{s}_1) A_{jk}(\mathbf{s}_2),$$

$\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d$. This type of approach can be useful if the observed multivariate process is linked in a varying way to some underlying and unobserved processes. Guhaniyogi et al. (2013) combined a low rank predictive process approach with the nonstationary LMC for computationally feasible modeling with large datasets.

The multivariate Matérn was extended to the nonstationary case by Kleiber and Nychka (2012). The basic idea is to allow the various Matérn parameters, variance, smoothness and length scale, to be spatially varying (Stein, 2005; Paciorek and Schervish, 2006), using normal scale mixtures (Schlather, 2010). For example, temperature fields exhibit longer range spatial dependence over the ocean than over land due to terrain driven nonstationarity, and a nonstationary Matérn

with spatially varying length scale parameter can capture this type of dependence without resorting to using disjoint models between ocean and land. In particular, the nonstationary multivariate Matérn supposes

$$\rho_{ii}(\mathbf{s}_1, \mathbf{s}_2) \propto \mathrm{M}(\mathbf{s}_1, \mathbf{s}_2 | \nu_i(\mathbf{s}_1, \mathbf{s}_2), a_i(\mathbf{s}_1, \mathbf{s}_2)),$$

$$\rho_{ij}(\mathbf{s}_1, \mathbf{s}_2) \propto \beta_{ij}(\mathbf{s}_1, \mathbf{s}_2) \mathrm{M}(\mathbf{s}_1, \mathbf{s}_2 | \nu_{ij}(\mathbf{s}_1, \mathbf{s}_2), a_{ij}(\mathbf{s}_1, \mathbf{s}_2)),$$

$\mathbf{s}_1, \mathbf{s}_2 \in \mathbb{R}^d$. An additional point here is that $\beta_{ij}(\mathbf{s}, \mathbf{s})$ is proportional to the collocated cross-correlation coefficient $\mathrm{cor}\{Z_i(\mathbf{s}), Z_j(\mathbf{s})\}$, that is, the strength of relationship between variables at the same location. This strength often varies spatially, for example minimum and maximum temperature are less correlated over highly mountainous regions than over plains where they exhibit greater dependence. Kleiber and Genton (2013) considered an approach to allowing this correlation coefficient to vary with location in such a way that it can be included with any arbitrary multivariate covariance choice, as long as each process has a nonzero nugget effect (which is not usually restrictive, as most processes exhibit small scale dependence that are typically modeled as nugget effects). Other authors have noted similar phenomena with other scientific data (Fuentes and Reich, 2013; Guhaniyogi et al., 2013).

Owing to the increasing complexity of nonstationary and multivariate models and the expertise required to decide on a framework as well as implement an estimation scheme, a few authors have considered nonparametric approaches to estimation. Extending Oehlert (1993) and Guillot, Senoussi and Monestiez (2001) to the multivariate case, Jun et al. (2011) and Kleiber, Katz and Rajagopalan (2013) worked with a nonparametric estimator of multivariate covariance that is free from model choice and is available throughout the observation domain. The underlying idea is to kernel smooth the empirical method-of-moments estimate of spatial covariance in a way that retains nonnegative definiteness and yields covariance estimates at any arbitrary location pairs, not only those with observations. Their nonparametric estimators are variations on the form

$$\hat{C}_{ij}(\mathbf{x}, \mathbf{y})$$

$$= \left(\sum_{k=1}^{n} \sum_{\ell=1}^{n} K_\lambda(\|\mathbf{x} - \mathbf{s}_k\|) \right.$$

$$(11) \qquad \cdot K_\lambda(\|\mathbf{y} - \mathbf{s}_\ell\|) Z_i(\mathbf{s}_k) Z_j(\mathbf{s}_\ell) \Big)$$

$$\cdot \left(\sum_{k=1}^{n} \sum_{\ell=1}^{n} K_\lambda(\|\mathbf{x} - \mathbf{s}_k\|) K_\lambda(\|\mathbf{y} - \mathbf{s}_\ell\|) \right)^{-1},$$

$\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, where $K_\lambda(r) = K(r/\lambda)$ is a positive kernel function with bandwidth λ. The displayed equation (11) is set up for the case when Z_i is mean zero for $i = 1, \ldots, p$, for instance representing residuals after a mean trend has been removed; the estimator can also be applied to centered residuals such as $Z_i(\mathbf{s}_k) - \bar{Z}_i$. This type of estimator can capture substantial nonstationarity that may be difficult to pick up parametrically (Kleiber, Katz and Rajagopalan, 2013). The nonparametric approach to estimation is primarily useful when replications of the multivariate random field are available. Although it can be applied when only a single field realization is available, we caution against its use given the well-known variability of empirical estimates in small samples.

The two methods of covariance and kernel convolution can also be extended to result in nonstationary matrix functions (Calder, 2007, 2008; Majumdar, Paul and Bautista, 2010). As with the univariate case, the convolution integrals are often intractable and must be estimated numerically, and parametric interpretations are sometimes ambiguous.

5. CROSS-COVARIANCE FUNCTIONS WITH SPECIAL FEATURES

5.1 Asymmetric Cross-Covariance Functions

All the stationary models described so far are symmetric, in the sense that $C_{ij}(\mathbf{h}) = C_{ji}(\mathbf{h})$, or equivalently, $C_{ij}(\mathbf{h}) = C_{ij}(-\mathbf{h})$. Although $C_{ij}(\mathbf{h}) = C_{ji}(-\mathbf{h})$ by definition, the aforementioned properties may not hold in general. Li, Genton and Sherman (2008) proposed a test of symmetry of the cross-covariance structure of multivariate random fields based on the asymptotic distribution of its empirical estimator. If the test rejects symmetry, then asymmetric cross-covariance functions are needed.

Li and Zhang (2011) proposed a general approach to render any stationary symmetric cross-covariance function asymmetric. The key idea is to notice that if $C_{ij}(\mathbf{h})$ is a valid symmetric cross-covariance function, then

$$(12) \quad C_{ij}^a(\mathbf{h}) = C_{ij}(\mathbf{h} + \mathbf{a}_i - \mathbf{a}_j), \quad \mathbf{h} \in \mathbb{R}^d,$$

is a valid asymmetric cross-covariance function for any vectors $\mathbf{a}_i \in \mathbb{R}^d$, $i = 1, \ldots, p$, such that $\mathbf{a}_i \neq \mathbf{a}_j$. Indeed, if $\mathbf{Z}(\mathbf{s}) = \{Z_1(\mathbf{s}), \ldots, Z_p(\mathbf{s})\}^T$ has cross-covariance functions $C_{ij}(\mathbf{h})$, then $\{Z_1(\mathbf{s} - \mathbf{a}_1), \ldots, Z_p(\mathbf{s} - \mathbf{a}_p)\}^T$ has cross-covariance functions $C_{ij}^a(\mathbf{h})$ given by (12), $i, j = 1, \ldots, p$. In particular, the construction (12) can be used to produce asymmetric versions of the LMC and the multivariate Matérn models.

The vectors $\mathbf{a}_1, \ldots, \mathbf{a}_p$ introduce delays that generate asymmetry in the cross-covariance structure. Because only the differences $\mathbf{a}_i - \mathbf{a}_j$ matter, one can impose a constraint such as $\mathbf{a}_1 + \cdots + \mathbf{a}_p = \mathbf{0}$ or $\mathbf{a}_1 = \mathbf{0}$ to ensure identifiability. Li and Zhang (2011) proposed to first estimate the marginal parameters of $C_{ij}^a(\mathbf{h})$ in (12), and then estimate the cross-parameters and $p - 1$ of the $\mathbf{a}_i$'s. Their simulations and data examples showed that asymmetric cross-covariance functions, when required, can achieve remarkable improvements in prediction over symmetric models. Apanasovich and Genton (2010) used a similar strategy to produce asymmetric spatio-temporal cross-covariance models based on latent dimensions; see Section 7.2. Inducing asymmetry in a nonstationary model is yet an open problem.

5.2 Compactly Supported Cross-Covariance Functions

Computational issues in the face of large datasets is a major problem in any spatial analysis, including likelihood calculations and/or co-kriging; see the review by Sun, Li and Genton (2012, Section 3.7). Especially, if the observation network is very large (even on the order of thousands), likelihood calculations and co-kriging equations are difficult or impossible to solve with standard covariance models, due to the dense unstructured observation covariance matrix. One approach to overcoming this difficulty is to induce sparsity in the covariance matrix, either by using a compactly supported covariance function as the model, or by covariance tapering, that is, multiplying a compactly supported nonnegative definite function against the model covariance (Furrer, Genton and Nychka, 2006; Kaufman, Schervish and Nychka, 2008). Then sparse matrix methods can be used to invert the covariance matrix, or find the determinant thereof.

Only recently have authors begun to consider this problem for multivariate random fields. Most of the currently available models are based on scale mixtures of the form

$$C_{ij}(\mathbf{h}) = \int (1 - \|\mathbf{h}\|/x)_+^v g_{ij}(x) \, \mathrm{d}x, \quad \mathbf{h} \in \mathbb{R}^d,$$

or variations on this theme (Reisert and Burkhardt, 2007; Porcu and Zastavnyi, 2011). Here, $v \geq (d + 1)/2$, and $\{g_{ij}(x)\}_{i,j=1}^p$ forms a valid cross-covariance matrix function. The generality of this construction gives rise to many interesting examples. For instance, with $g_{ij}(x) = x^v(1 - x/b)_+^{\gamma_{ij}}$ where $\gamma_{ij} = (\gamma_i + \gamma_j)/2$ and $\gamma_i > 0$ for all $i = 1, \ldots, p$ we have the multivariate

Askey taper

$$C_{ij}(\mathbf{h}) = b^{\nu+1} B(\gamma_{ij} + 1, \nu + 1) \left(1 - \frac{\|\mathbf{h}\|}{b}\right)^{\nu+\gamma_{ij}+1},$$

$\|\mathbf{h}\| < b$, and 0 otherwise, where B is the beta function (Porcu et al., 2013). Kleiber and Porcu (2015) provided a nonstationary extension of this model, while Porcu et al. (2013) considered similar ideas for Buhmann functions and B-splines. Daley, Porcu and Bevilacqua (2015) obtained multivariate Askey functions with different compact supports b_{ij} and the multivariate analogue of Wendland functions. The latter provide a tool for tapering cross-covariance functions such as the multivariate Matérn. Recent results on equivalence of Gaussian measures of multivariate random fields by Ruiz-Medina and Porcu (2015) will allow for assessing the statistical properties of multivariate tapers. Du and Ma (2013) derived compactly supported classes of the Pólya type. Although there has been a flurry of recent activity, much additional work remains in implementing these models in real world applications, exploring covariance tapering and understanding limitations of stationary constructions.

5.3 Cross-Covariance Functions on the Sphere

Many multivariate datasets from environmental and climate sciences are collected over large portions of the Earth, for example, by satellites and, therefore, cross-covariance functions on the sphere $\mathbb{S}^2$ in $\mathbb{R}^3$ are in need. Consider a multivariate process on the sphere for which the ith variable is described by $Z_i(L, l)$, $i = 1, \ldots, p$, with L denoting latitude and l denoting longitude. Jun (2011) constructed cross-covariance functions by applying differential operators with respect to latitude and longitude to the process on the sphere. Furthermore, Jun (2011) studied nonstationary models of cross-covariances with respect to latitude, so-called axially symmetric, and longitudinally irreversible cross-covariance functions for which

$$\text{cov}\{Z_i(L_1, l_1), Z_j(L_2, l_2)\}$$
$$\neq \text{cov}\{Z_i(L_1, l_2), Z_j(L_2, l_1)\},$$

$(L_1, l_1) \in \mathbb{S}^2$, $(L_2, l_2) \in \mathbb{S}^2$. All the models described in Jun (2011) are valid for the chordal distance, that is, the Euclidean distance in $\mathbb{R}^3$ between points on $\mathbb{S}^2$. Castruccio and Genton (2014) relaxed the assumption of axial symmetry for univariate random fields on the sphere and the extension of their work to multivariate random fields on the sphere remains an open problem. Gneiting (2013) provided a very thorough

study of positive definite functions on a sphere that can be used as covariances. Du, Ma and Li (2013) developed a characterization of isotropic and continuous variogram matrix functions on the sphere, extending some of the ideas of Ma (2012) who characterized continuous and isotropic covariance matrix functions on the sphere using Gegenbauer polynomials. Because the great circles are the geodesics on the sphere, they are the natural metric to measure distances in this context. Porcu, Bevilacqua and Genton (2014) developed cross-covariance functions of the great circle distances on the sphere. In particular, they studied multivariate Matérn models as functions of the great circle distance on the sphere. Recently, Jun (2014) developed nonstationary Matérn cross-covariance models whose smoothness parameters vary over space and with large-scale nonstationarity obtained with the aforementioned differential operators.

6. DATA EXAMPLES

We illustrate a selection of the above cross-covariance models on two data examples. First, a set of reanalysis climate model output that represents spatially gridded data. Second, a set of observational temperature data that illustrates spatially irregularly located data.

6.1 Climate Model Output Data

The specific reanalysis dataset in use is a National Centers for Environmental Protection-driven (NCEP) run of the updated Experimental Climate Prediction Center (ECP2) model, which was originally run as part of the North American Regional Climate Change Assessment Program (NARCCAP) climate modeling experiment (Mearns et al., 2009). Reanalysis data can be thought of as an estimate of the true state of the atmosphere for a given period. The variables we use are average summer temperature and cube-root precipitation (summer being comprised of June, July and August; JJA) over a region of the midwest United States that is largely an agricultural region with relatively constant terrain. The cube-root transformation reduces skewness in the precipitation output and brings the distribution closer to Gaussian. For each grid cell, we calculate a pointwise spatially varying mean as the arithmetic average of all 24 years of model output from 1981 through 2004. The data considered then are 24 years of residuals, having removed this spatially varying mean from each year's reanalysis output for the two variables of temperature and cube-root precipitation.

The residuals are assumed to be independent between years, and are additionally assumed to be realizations from a mean zero bivariate Gaussian process (both assumptions are supported by exploratory analysis).

Figure 1 contains an example set of reanalysis residuals for the year 1989. By eye, it appears that temperature residuals are smoother over space, while precipitation is apparently rougher, while both seem to have similar correlation length scales. The two variables are strongly negatively correlated, with an empirical correlation coefficient of -0.67. This situation, with negative and strong cross-correlation and both variables exhibiting distinct levels of smoothness, provides numerous challenges to available cross-correlation models. Call $T(\mathbf{s}, t)$ and $P(\mathbf{s}, t)$ the temperature and precipitation residual at location $\mathbf{s}$ in year t, respectively (recalling that, although indexed by year, the processes are viewed as temporally-independent).

Of the above models, we compare six to an independence assumption, that is, where temperature and precipitation residuals are assumed to be independent; for the independence model, each variable is assumed to follow a Matérn covariance, and parameters are estimated by maximum likelihood. The first nontrivial bivariate model is the parsimonious Matérn, whose parameters we estimate by maximum likelihood. The second model is a nearly full bivariate Matérn, where we set the cross-covariance smoothness ν_{TP}, T representing temperature and P precipitation, to be the arithmetic average of the marginal smoothnesses. For the full bivariate Matérn, we set marginal parameters to be those of the independence model, and conditional on these, estimate the remaining cross-covariance length scale a_{TP} and cross-correlation coefficient ρ_{TP} by maximum likelihood. We additionally consider two variations on the bivariate parsimonious Matérn, one using a lagged covariance of Li and Zhang (2011) (see Section 5.1), and a nonstationary Matérn with spatially varying variances for both variables. Spatially varying variances are estimated empirically at each grid cell, and conditional on these, the remaining parameters are estimated by maximum likelihood. We also consider a linear model of coregionalization,

$$T(\mathbf{s}, t) = a_{11} Z_1(\mathbf{s}, t),$$

$$P(\mathbf{s}, t) = a_{12} Z_1(\mathbf{s}, t) + a_{22} Z_2(\mathbf{s}, t),$$

where Z_1 and Z_2 are independent mean zero spatial processes with Matérn covariances. We opt for this formulation since temperature is expected to be smoother than precipitation, and our goal is to preserve this feature within the statistical model. Parameters are estimated by maximum likelihood. Finally, we additionally consider two latent dimensional models. The first is parameterized by (9), except without a nugget effect, and the second is built via

$$T(\mathbf{s}, t) = b_{11} Z(\mathbf{s}, t) + b_{12} Z_1(\mathbf{s}, t),$$

$$P(\mathbf{s}, t) = b_{21} Z(\mathbf{s}, t) + b_{22} Z_2(\mathbf{s}, t),$$

where $Z(\mathbf{s}, t)$ has a latent dimensional covariance of the form

$$C(\mathbf{h}) = \frac{1}{(\|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\| + 1)^\beta} \exp\left\{ \frac{-\alpha \|\mathbf{h}\|^2}{(\|\boldsymbol{\xi}_i - \boldsymbol{\xi}_j\| + 1)^\beta} \right\},$$

$\mathbf{h} \in \mathbb{R}^2$, and Z_1, Z_2 are independent with Matérn correlations. This choice for Z allows the temperature process to retain smoother behavior at the origin than precipitation, whereas the model of (9) forces exponential-like behavior at the origin.

Table 1 contains the parsimonious and full bivariate Matérn parameter estimates. Note the smoothness parameter of the temperature field is approximately 1.3,
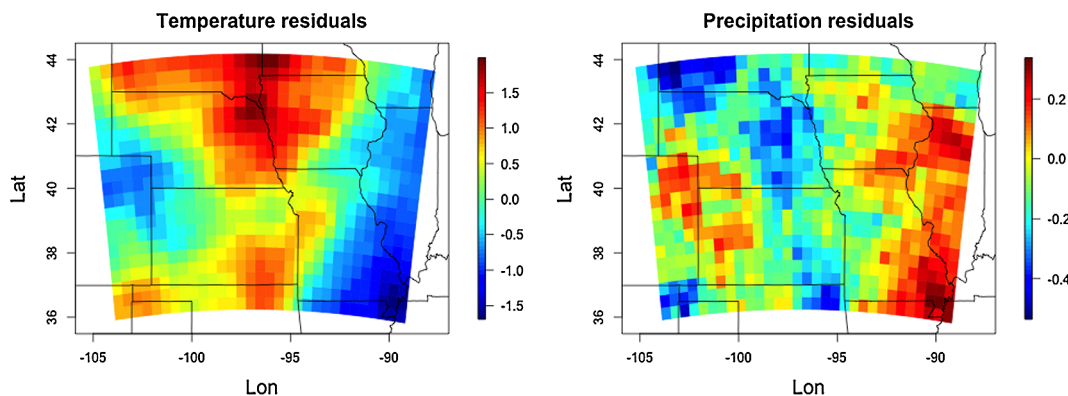

FIG. 1. *Example residuals from 1989 after removing a spatially varying mean from NCEP-driven ECP2 regional climate model run for the variables of average summer temperature and precipitation. Units are degrees Celsius for temperature and centimeters for precipitation.*

TABLE 1

Maximum likelihood estimates of parameters for full and parsimonious bivariate Matérn models, applied to the NARCCAP model data.
Units are degrees Celsius for temperature, centimeters for precipitation, and kilometers for distances

Model	σ_T	σ_P	ν_T	ν_P	$1/a_T$	$1/a_P$	$1/a_{TP}$	ρ_{TP}
Full	1.63	0.19	1.31	0.55	384.3	361.6	420.1	−0.60
Parsimonious	1.61	0.19	1.33	0.54	367.1	–	–	−0.49

indicating a relatively smooth field, which supports the theoretical analysis of North, Wang and Genton (2011); on the other hand, precipitation has a smoothness of approximately 0.55, suggesting an exponential model may work well. Both variables have similar length scale parameters, which suggests the assumptions of the parsimonious Matérn model may be reasonable for this particular dataset. The cross-correlation coefficient is estimated to be strongly negative in both cases, with the full Matérn slightly closer to the empirical cross-correlation.

Table 2 contains log likelihood values for the various models considered. Evidently, the parsimonious, full and parsimonious lagged Matérn all have likelihood values on the same order, which are all superior to the LMC, independent Matérn and latent dimensional models. We remark that, given the smooth nature of the temperature field, the latent dimensional model of (9) is not expected to perform as well, as it fixes the smoothness of the temperature field at $\nu = 0.5$, while on the other hand the latent dimensional model using a shared process with squared exponential covariance performs nearly as well as the Matérn alternatives. The nonstationary extension of the parsimonious Matérn exhibits the largest log likelihood, improving the next best model by over 1000. This suggests that the bivariate field indeed exhibits nonstationarity, and there

may be other modeling improvements that can be explored with new nonstationary cross-covariance developments.

Finally, we perform a small pseudo cross-validation study. We hold out the bivariate model output at a randomly chosen 90% of spatial locations consistent over all time points. We then co-krige the remaining 10% (62 locations) to the held out grid cells using parameter estimates based on the entire dataset. As the residual process is assumed to be independent between years, co-kriging is performed separately for each year. Root mean squared error (RMSE) and the continuous ranked probability score (CRPS) are used to validate interpolation quality, averaged over all held out locations and years. We repeat this experiment ten times for different randomly chosen sets of held out spatial locations and average the resulting scores; the results are displayed in Table 2. Generally speaking, all models are effectively equivalent in terms of predictive ability, except for the nonstationary extension to the parsimonious Matérn, which appears to improve both predictive quantities for temperature especially. Perhaps surprisingly, the independent Matérn performs as well for interpolation, although this has not been the case with all datasets (Gneiting, Kleiber and Schlather, 2010).

TABLE 2

Comparison of log likelihood values and pseudo cross-validation scores averaged over ten cross-validation replications for various
multivariate models on the NARCCAP model data residuals for temperature (T) and precipitation (P)

	Log likelihood	RMSE (T)	CRPS (T)	RMSE (P)	CRPS (P)
Nonstationary parsimonious Matérn	53564.5	0.168	0.084	0.085	0.047
Parsimonious lagged Matérn	52563.7	0.179	0.090	0.087	0.048
Full Matérn	52560.1	0.178	0.090	0.087	0.048
Parsimonious Matérn	52556.9	0.179	0.090	0.087	0.048
Latent dimension	52028.8	0.180	0.091	0.088	0.049
LMC	51937.0	0.179	0.091	0.090	0.050
Independent Matérn	50354.5	0.180	0.091	0.088	0.049
Latent dimension of (9)	48086.3	0.195	0.100	0.088	0.048

Comparison of log likelihood values and pseudo cross-validation scores averaged over 100 *cross-validation replications for various multivariate models on the USHCN observed temperature residuals for maximum temperature (max) and minimum temperature (min)*

	Log likelihood	RMSE (min)	CRPS (min)	RMSE (max)	CRPS (max)
Parsimonious lagged Matérn	−414.0	3.18	1.83	3.14	1.79
Parsimonious Matérn	−414.9	3.22	1.85	3.16	1.80
LMC	−415.7	3.22	1.85	3.16	1.80
Latent dimension	−416.2	3.23	1.86	3.18	1.81
Latent dimension of (9)	−419.1	3.24	1.86	3.17	1.81
Independent Matérn	−427.6	3.41	1.94	3.35	1.91

6.2 Observational Temperature Data

The second example we consider is a bivariate minimum and maximum temperature observational dataset. Observations are available at stations that are part of the United States Historical Climatology Network (Peterson and Vose, 1997) over the state of Colorado. Stations in the USHCN form the highest quality observational climate network in the United States; observations are subject to rigorous quality control.

We consider bivariate daily temperature residuals (i.e., having removed the state-wide mean) on September 19, 2004, a day which has good network coverage with observations being available at 94 stations. Exploratory Q–Q plots suggest the residuals are well modeled marginally as Gaussian processes; we suppose the bivariate process is a realization from a bivariate Gaussian process with zero mean.

We entertain the same set of bivariate models as in the previous example subsection. Due to the fact that the data are observational, we augment each process' covariance with a nugget effect. We begin by estimating the independent Matérn model separately for both minimum and maximum temperature residuals by maximum likelihood. Since the nugget effect is tied to marginal process behavior, we fix the estimated nugget effects at their marginal estimates, and estimate all other covariance parameters from the remaining bivariate models by maximum likelihood, conditional on these marginal nugget estimates. We remove both the bivariate Matérn and nonstationary model from consideration, as these are both difficult to estimate given a single realization of the spatial process.

On top of comparing in sample log likelihood values, we additionally consider a pseudo cross-validation study, leaving out a randomly selected 25% of locations, and co-krige the remaining bivariate observations to these held out locations. This pseudo cross-validation procedure is repeated 100 times, and Table 3 contains the averaged scores from this study.

Contrasting with the results of the NARCCAP example, we now see the predictive benefit of considering multivariate second-order structures. Generally, predictive RMSE and CRPS are improved by between 6–7% when co-kriging using the parsimonious lagged Matérn, as compared to marginally kriging each variable. A potential explanation for the improvement here as compared to the NARCCAP example is that in the current study, the observations are subject to measurement error, and thus the greater uncertainty in estimating the bivariate surface is more readily quantified using an appropriate bivariate covariance model.

7. DISCUSSION

7.1 Specialized Cross-Covariance Functions

The models introduced so far cover the broad majority of usual datasets requiring multivariate models. However, specialized scenarios sometimes arise, and call for novel developments. For instance, some constructions involve modeling variables that exhibit long range dependence. Ma (2011c) examined a construction for all variables having long or short range dependence utilizing univariate variograms; and Ma (2011a) explored the relationship between multivariate covariances and variograms. Kleiber and Porcu (2015) derived a nonstationary construction that allows individual variables to be a spatially varying mixture of short and long range dependence, as well as having substantial cross-correlation between variables (with possibly opposing short/long range dependence); their construction is a special case of a multivariate generalization of the univariate Cauchy class of covariance (Gneiting and Schlather, 2004). Hristopoulos and Porcu (2014) defined the multivariate analogue of Spartan Gibbs random fields, obtained through using Hamiltonian functionals.

Ma (2011b) also studied various approaches to produce valid cross-covariance functions based on differentiation of univariate covariance functions and on scale mixtures of covariance matrix functions. Alternatively, Ma (2011d) provided constructions of variogram matrix functions, and Du and Ma (2012) introduced an approach to building variogram matrix functions based on a univariate variogram model.

We close this section by pointing out a recent novel approach to generating valid matrix covariances by considering stochastic partial differential equations (SPDEs); Hu et al. (2013) used systems of SPDEs to simultaneously model temperature and humidity, yielding computationally efficient means to analysis by approximating a Gaussian random field by a Gaussian Markov random field.

7.2 Spatio-Temporal Cross-Covariance Functions

So far, the cross-covariance models that we described were aimed at spatial multivariate random fields. When adding the time dimension, the resulting spatio-temporal multivariate random field, $\mathbf{Z}(\mathbf{s}, t)$, has stationary cross-covariance functions $C_{ij}(\mathbf{h}, u)$, where u denotes a time lag. All the previous spatial cross-covariance models can be straightforwardly extended to the spatio-temporal setting, for example, Rouhani and Wackernagel (1990), Choi et al. (2009), Berrocal, Gelfand and Holland (2010) and De Iaco et al. (2013), De Iaco, Palma and Posa (2013) developed space–time versions of the linear model of coregionalization. Gelfand, Banerjee and Gamerman (2005) used a dynamic approach for multivariate space–time data using coregionalization.

Based on the concept of latent dimensions described in Section 2.3, Apanasovich and Genton (2010) have extended a class of spatio-temporal covariance functions for univariate random fields due to Gneiting (2002) to the multivariate setting. Specifically, if $\varphi_1(t)$, $t \geq 0$, is a completely monotone function and $\psi_1(t)$, $\psi_2(t)$, $t \geq 0$, are positive functions with completely monotone derivatives, then

$$
(13) \quad \begin{aligned} C(\mathbf{h}, u, \mathbf{v}) &= \frac{\sigma^2}{[\psi_1\{u^2/\psi_2(\|\mathbf{v}\|^2)\}]^{d/2}\{\psi_2(\|\mathbf{v}\|^2)\}^{1/2}} \\ &\quad \cdot \varphi_1\left[\frac{\|\mathbf{h}\|^2}{\psi_1\{u^2/\psi_2(\|\mathbf{v}\|^2)\}}\right], \end{aligned}
$$

is a valid stationary covariance function on $\mathbb{R}^{d+1+k}$ that can be used to model cross-covariance functions with $\mathbf{v} = \boldsymbol{\xi}_i - \boldsymbol{\xi}_j$. When $\psi_2(t) \equiv 1$, Gneiting's class

is retrieved. The case $\mathbf{v} = \mathbf{0}$ yields a common spatio-temporal covariance function for each variable that can be made different through a LMC-type construction. Also judicious choices of the functions in (13) allow one to control nonseparability between space and time, between space and variables, and between time and variables; see Apanasovich and Genton (2010) for various illustrative examples.

To further introduce asymmetry in spatio-temporal cross-covariance functions, Apanasovich and Genton (2010) have proposed two approaches based on latent dimensions. Using the notation of Section 2.3, the first type of asymmetric spatio-temporal cross-covariance is

$$
(14) \quad C_{ij}^a(\mathbf{h}, u) = C(\mathbf{h}, u - \boldsymbol{\lambda}_\xi^{\mathsf{T}}(\boldsymbol{\xi}_i - \boldsymbol{\xi}_j), \boldsymbol{\xi}_i - \boldsymbol{\xi}_j),
$$

$\mathbf{h} \in \mathbb{R}^d, u \in \mathbb{R}$, where C is a valid covariance function on $\mathbb{R}^{d+k}$ of a univariate random field and $\boldsymbol{\lambda}_\xi \in \mathbb{R}^k$, $1 \leq k \leq p$, controls the delay in time that creates asymmetry. There is no time delay if and only if $\boldsymbol{\lambda}_\xi = \mathbf{0}$ or $i = j$. The second type of asymmetric spatio-temporal cross-covariance is

$$
(15) \quad C_{ij}^a(\mathbf{h}, u) = C(\mathbf{h} - \boldsymbol{\gamma}_h u, u, \boldsymbol{\xi}_i - \boldsymbol{\xi}_j - \boldsymbol{\gamma}_\xi u),
$$

$\mathbf{h} \in \mathbb{R}^d, u \in \mathbb{R}$, where the velocity vectors $\boldsymbol{\gamma}_h \in \mathbb{R}^d$ and $\boldsymbol{\gamma}_\xi \in \mathbb{R}^k$ are responsible for the lack of symmetry. When $u \neq 0$, this model is spatially anisotropic. Combinations of models (14) and (15) are possible.

7.3 Physics-Constrained Cross-Covariance Functions

Especially for geophysical processes, often there are physical constraints on a system of variables that must be obeyed by any stochastic model. For instance, Buell (1972) explored valid covariance models for geostrophic wind that must satisfy physical relationships for isotropic geophysical flow including geopotential, longitudinal wind components and transverse wind components.

In a similar vein, a number of physical processes, especially in fluid dynamics, involve fields with specialized restrictions such as being divergence free. Scheuerer and Schlather (2012) developed matrix-valued covariance functions for divergence-free and curl-free random vector fields, which are based on combinations of derivatives of a specified variogram and extend earlier work by Narcowich and Ward (1994).

Constantinescu and Anitescu (2013) introduced a framework for building valid matrix-valued covariance functions when the constituent processes have known

physical constraints relating their behavior. By approximating a nonlinear physical relationship between variables through series expansions and closures, the authors develop physically-based matrix covariance classes. They explored large-scale geostrophic wind as a case study, and illustrated that physically motivated cross-correlation models can substantially outperform independence models.

North, Wang and Genton (2011) studied spatio-temporal correlations for temperature fields arising from simple energy-balance climate models, that is, white-noise-driven damped diffusion equations. The resulting spatial correlation on the plane is of Matérn type with smoothness parameter $\nu = 1$, although rougher temperature fields are expected due to terrain irregularities for example. Derivations for temperature fields on a uniform sphere were presented as well. Whether these results can be extended to other variables such as pressure and wind fields, and possibly lead to Matérn cross-covariance models of type (10), is an open question.

7.4 Open Problems

Finally, there are many open problems that call for more research. The most fundamental question is the theoretical characterization of the allowable classes of multivariate covariances. For instance, given two marginal covariances, what is the valid class of possible cross-covariances that still results in a nonnegative definite structure? Such a characterization is an unsolved problem. Additional to characterization, the companion theoretical question is the utility of cross-covariance models. Given the two data examples in this review, a natural question is: for the purposes of co-kriging, in what situations are the use of nontrivial cross-covariances beneficial? Although it is traditional to focus on kriging and co-kriging in the geostatistical literature, we wish to additionally emphasize the utility of these models for simulation of multivariate random fields. Indeed, without flexible cross-covariance models, it is impossible to simulate multiple fields with nontrivial dependencies.

The power exponential class of covariances is a useful marginal class of covariances, but to the best of our knowledge, a characterization of parameters for the validity of the multivariate version

$$\rho_{ij}(\mathbf{h}) = \beta_{ij} \exp\left\{ -\left(\frac{\|\mathbf{h}\|}{\phi_{ij}} \right)^{\kappa_{ij}} \right\}, \quad \mathbf{h} \in \mathbb{R}^d,$$

is not known. Although we believe that the multivariate Matérn model (10) has more flexibility, this is still

an interesting question, especially as this set of covariances requires no calculations involving Bessel functions.

The extension of spatial extremes to the case of multiple variables has not been explored yet except for the recent proposal of Genton, Padoan and Sang (2015) who considered multivariate max-stable spatial processes. The aim of that research is to describe the behavior of extreme events of several variables across space, such as extreme rainfall and extreme temperature for example. This requires flexible and physically-realistic cross-covariance models and therefore the families described herein may play an important role for such applications.

Recently, there has been some new interest in other types of random fields than the usual Gaussian case. Mittag–Leffler fields contain the Gaussian case as a subset, but are specified in terms of an infinite series expansion that is unwieldy for applications (Ma, 2013b). Another option is a multivariate extension of the Student's t distribution, a t-vector distribution (Ma, 2013a); these seem to be more promising for applications, and some exploration of the utility of these types of models is called for. Finally, hyperbolic vector random fields contain the Student's t as a limiting case, although model interpretation, estimation and implementation remain unexplored (Du et al., 2012).

There is also a need for valid multivariate cross-covariance functions for spatial data on a lattice. Although one can apply any of the models mentioned in this manuscript to lattice data, the extension of univariate Markov random field models is another route. For instance, Gelfand and Vounatsou (2003) have studied proper multivariate conditional autoregressive models. Daniels, Zhou and Zou (2006) proposed a class of conditionally specified space–time models for multivariate processes geared to situations where there is a sparse spatial coverage of one of the processes and a much more dense coverage of the other processes. This is motivated by an application to particulate matter and ozone data. Sain and Cressie (2007) also developed Markov random field models for multivariate lattice data.

Many additional open questions remain, including theoretical development of estimation in the multivariate context (Pascual and Zhang, 2006). Vargas-Guzmán, Warrick and Myers (1999) looked at the relationship between support size and relationship between variables, but relatively few have explored this phenomenon in the multivariate case. Finally, there is a need to better understand and explore

the intimate connection between multivariate spline smoothers, co-kriging and multivariate numerical analysis (Beatson, zu Castell and Schrödl, 2011; Fuselier, 2008; Narcowich and Ward, 1994; Reisert and Burkhardt, 2007).

REFERENCES

ALMEIDA, A. S. and JOURNEL, A. G. (1994). Joint simulation of multiple variables with a Markov-type coregionalization model. *Math. Geol.* **26** 565–588.

ALONSO-MALAVER, C., PORCU, E. and GIRALDO, R. (2013). Multivariate versions of walks through dimensions and Schoenberg measures. Technical report, Univ. Tecnica Federico Santa Maria, Valparaiso, Chile.

ALONSO-MALAVER, C., PORCU, E. and GIRALDO, R. (2015). Multivariate and multiradial Schoenberg measures with their dimension walk. *J. Multivariate Anal.* **133** 251–265. MR3282029

ÁLVAREZ, M. A., ROSASCO, L. and LAWRENCE, N. D. (2012). Kernels for vector-valued functions: A review. *Found. Trends Mach. Learn.* **3** 195–266.

APANASOVICH, T. V. and GENTON, M. G. (2010). Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika* **97** 15–30. MR2594414

APANASOVICH, T. V., GENTON, M. G. and SUN, Y. (2012). A valid Matérn class of cross-covariance functions for multivariate random fields with any number of components. *J. Amer. Statist. Assoc.* **107** 180–193. MR2949350

BEATSON, R. K., ZU CASTELL, W. and SCHRÖDL, S. J. (2011). Kernel-based methods for vector-valued data with correlated components. *SIAM J. Sci. Comput.* **33** 1975–1995. MR2831042

BERROCAL, V. J., GELFAND, A. E. and HOLLAND, D. M. (2010). A bivariate space–time downscaler under space and time misalignment. *Ann. Appl. Stat.* **4** 1942–1975. MR2829942

BHAT, K., HARAN, M. and GOES, M. (2010). Computer model calibration with multivariate spatial output: A case study. In *Frontiers of Statistical Decision Making and Bayesian Analysis* (M.-H. Chen, D. K. Dey, P. Müller, D. Sun and K. Ye, eds.) 168–184. Springer, New York. MR2766461

BOCHNER, S. (1955). *Harmonic Analysis and the Theory of Probability*. Univ. California Press, Berkeley and Los Angeles. MR0072370

BORNN, L., SHADDICK, G. and ZIDEK, J. V. (2012). Modeling nonstationary processes through dimension expansion. *J. Amer. Statist. Assoc.* **107** 281–289. MR2949359

BOURGAULT, G. and MARCOTTE, D. (1991). Multivariable variogram and its application to the linear model of coregionalization. *Math. Geol.* **23** 899–928.

BROWN, P. J., LE, N. D. and ZIDEK, J. V. (1994). Multivariate spatial interpolation and exposure to air pollutants. *Canad. J. Statist.* **22** 489–509. MR1321471

BUELL, C. E. (1972). Correlation functions for wind and geopotential on isobaric surfaces. *J. Appl. Meteorol.* **11** 51–59.

CALDER, C. A. (2007). Dynamic factor process convolution models for multivariate space–time data with application to air quality assessment. *Environ. Ecol. Stat.* **14** 229–247. MR2405328

CALDER, C. A. (2008). A dynamic process convolution approach to modeling ambient particulate matter concentrations. *Environmetrics* **19** 39–48. MR2416543

CASTRUCCIO, S. and GENTON, M. G. (2014). Beyond axial symmetry: An improved class of models for global data. *Stat* **3** 48–55.

CHOI, J., REICH, B. J., FUENTES, M. and DAVIS, J. M. (2009). Multivariate spatial–temporal modeling and prediction of speciated fine particles. *J. Stat. Theory Pract.* **3** 407–418. MR2751608

CONSTANTINESCU, E. M. and ANITESCU, M. (2013). Physics-based covariance models for Gaussian processes with multiple outputs. *Int. J. Uncertain. Quantif.* **3** 47–71. MR3044910

CRAMÉR, H. (1940). On the theory of stationary random processes. *Ann. of Math.* (2) **41** 215–230. MR0000920

CRESSIE, N. A. C. (1993). *Statistics for Spatial Data. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics*. Wiley, New York. MR1239641

CRESSIE, N. and WIKLE, C. K. (1998). The variance-based cross-variogram: You can add apples and oranges. *Math. Geol.* **30** 789–799. MR1646242

DALEY, D. J., PORCU, E. and BEVILACQUA, M. (2015). Classes of compactly supported correlation functions for multivariate random fields. *Stoch. Environ. Risk Assess.* To appear.

DANIELS, M. J., ZHOU, Z. and ZOU, H. (2006). Conditionally specified space–time models for multivariate processes. *J. Comput. Graph. Statist.* **15** 157–177. MR2269367

DE IACO, S., PALMA, M. and POSA, D. (2013). Prediction of particle pollution through spatio-temporal multivariate geostatistical analysis: Spatial special issue. *Adv. Stat. Anal.* **97** 133–150. MR3045764

DE IACO, S., MYERS, D. E., PALMA, M. and POSA, D. (2013). Using simultaneous diagonalization to identify a space–time linear coregionalization model. *Math. Geosci.* **45** 69–86. MR3018750

DU, J. and MA, C. (2012). Variogram matrix functions for vector random fields with second-order increments. *Math. Geosci.* **44** 411–425. MR2925032

DU, J. and MA, C. (2013). Vector random fields with compactly supported covariance matrix functions. *J. Statist. Plann. Inference* **143** 457–467. MR2995107

DU, J., MA, C. and LI, Y. (2013). Isotropic variogram matrix functions on spheres. *Math. Geosci.* **45** 341–357. MR3107159

DU, J., LEONENKO, N., MA, C. and SHU, H. (2012). Hyperbolic vector random fields with hyperbolic direct and cross covariance functions. *Stoch. Anal. Appl.* **30** 662–674. MR2946043

FANSHAWE, T. R. and DIGGLE, P. J. (2012). Bivariate geostatistical modelling: A review and an application to spatial variation in radon concentrations. *Environ. Ecol. Stat.* **19** 139–160. MR2946168

FUENTES, M. (2002). Spectral methods for nonstationary spatial processes. *Biometrika* **89** 197–210. MR1888368

FUENTES, M. and REICH, B. (2013). Multivariate spatial nonparametric modelling via kernel processes mixing. *Statist. Sinica* **23** 75–97. MR3076159

FURRER, R. (2005). Covariance estimation under spatial dependence. *J. Multivariate Anal.* **94** 366–381. MR2167920

FURRER, R. and GENTON, M. G. (2011). Aggregation-cokriging for highly multivariate spatial data. *Biometrika* **98** 615–631. MR2836410

FURRER, R., GENTON, M. G. and NYCHKA, D. (2006). Covariance tapering for interpolation of large spatial datasets. *J. Comput. Graph. Statist.* **15** 502–523. MR2291261

FUSELIER, E. J. (2008). Improved stability estimates and a characterization of the native space for matrix-valued RBFs. *Adv. Comput. Math.* **29** 269–290. MR2438345

GASPARI, G. and COHN, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quarterly J. Roy. Meteorol. Soc.* **125** 723–757.

GASPARI, G., COHN, S. E., GUO, J. and PAWSON, S. (2006). Construction and application of covariance functions with variable length-fields. *Quarterly J. Roy. Meteorol. Soc.* **132** 1815–1838.

GELFAND, A. E., BANERJEE, S. and GAMERMAN, D. (2005). Spatial process modelling for univariate and multivariate dynamic spatial data. *Environmetrics* **16** 465–479. MR2147537

GELFAND, A. E. and VOUNATSOU, P. (2003). Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics* **4** 11–25.

GELFAND, A. E., SCHMIDT, A. M., BANERJEE, S. and SIRMANS, C. F. (2004). Nonstationary multivariate process modeling through spatially varying coregionalization. *Test* **13** 263–312. MR2154003

GENTON, M. G., PADOAN, S. and SANG, H. (2015). Multivariate max-stable spatial processes. *Biometrika* **102** 215–230. MR3335107

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

GNEITING, T. (2013). Strictly and non-strictly positive definite functions on spheres. *Bernoulli* **19** 1327–1349. MR3102554

GNEITING, T., GENTON, M. G. and GUTTORP, P. (2007). Geostatistical space-time models, stationarity, separability and full symmetry. In *Statistics of Spatio-Temporal Systems. Monographs in Statistics and Applied Probability* (B. Finkenstaedt, L. Held and V. Isham, eds.) 151–175. Chapman & Hall/CRC Press, Boca Raton, FL.

GNEITING, T. and GUTTORP, P. (2006). Studies in the history of probability and statistics. XLIX. On the Matérn correlation family. *Biometrika* **93** 989–995. MR2285084

GNEITING, T., KLEIBER, W. and SCHLATHER, M. (2010). Matérn cross-covariance functions for multivariate random fields. *J. Amer. Statist. Assoc.* **105** 1167–1177. MR2752612

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

GOFF, J. A. and JORDAN, T. H. (1988). Stochastic modeling of seafloor morphology: inversion of sea beam data for second-order statistics. *J. Geophys. Res.* **93** 13589–13608.

GOULARD, M. and VOLTZ, M. (1992). Linear coregionalization model: Tools for estimation and choice of cross-variogram matrix. *Math. Geol.* **24** 269–282.

GRZEBYK, M. and WACKERNAGEL, H. (1994). Multivariate analysis and spatial/temporal scales: Real and complex models. In *Proceedings of XVIIth International Biometric Conference, Hamilton, Ontario, Canada* **1** 19–33.

GUHANIYOGI, R., FINLEY, A. O., BANERJEE, S. and KOBE, R. K. (2013). Modeling complex spatial dependencies: Low-rank spatially varying cross-covariances with application to soil nutrient data. *J. Agric. Biol. Environ. Stat.* **18** 274–298. MR3110894

GUILLOT, G., SENOUSSI, R. and MONESTIEZ, P. (2001). A positive definite estimator of the non stationary covariance of random fields. In *GeoENV 2000: Third European Conference on Geostatistics for Environmental Applications* (P. Monestiez, D. Allard and R. Froidevaux, eds.) 333–344. Kluwer Academic, Dordrecht.

HELTERBRAND, J. D. and CRESSIE, N. (1994). Universal cokriging under intrinsic coregionalization. *Math. Geol.* **26** 205–226. MR1267008

HRISTOPOULOS, D. and PORCU, E. (2014). Vector Spartan spatial random field models. *Probabilistic Engineering Mechanics* **37** 84–92.

HU, X., STEINSLAND, I., SIMPSON, D., MARTINO, S. and RUE, H. (2013). Spatial modelling of temperature and humidity using systems of stochastic partial differential equations. Available at arXiv:1307.1402v1.

HUANG, C., YAO, Y., CRESSIE, N. and HSING, T. (2009). Multivariate intrinsic random functions for cokriging. *Math. Geosci.* **41** 887–904. MR2557237

JOURNEL, A. G. (1999). Markov models for cross-covariances. *Math. Geol.* **31** 955–964. MR1717033

JUN, M. (2011). Non-stationary cross-covariance models for multivariate processes on a globe. *Scand. J. Stat.* **38** 726–747. MR2859747

JUN, M. (2014). Matérn-based nonstationary cross-covariance models for global processes. *J. Multivariate Anal.* **128** 134–146. MR3199833

JUN, M. and GENTON, M. G. (2012). A test for stationarity of spatio-temporal random fields on planar and spherical domains. *Statist. Sinica* **22** 1737–1764. MR3027105

JUN, M., SZUNYOGH, I., GENTON, M. G., ZHANG, F. and BISHOP, C. H. (2011). A statistical investigation of the sensitivity of ensemble-based Kalman filters to covariance filtering. *Mon. Weather Rev.* **139** 3036–3051.

KAUFMAN, C. G., SCHERVISH, M. J. and NYCHKA, D. W. (2008). Covariance tapering for likelihood-based estimation in large spatial data sets. *J. Amer. Statist. Assoc.* **103** 1545–1555. MR2504203

KLEIBER, W. and GENTON, M. G. (2013). Spatially varying cross-correlation coefficients in the presence of nugget effects. *Biometrika* **100** 213–220. MR3034334

KLEIBER, W., KATZ, R. W. and RAJAGOPALAN, B. (2013). Daily minimum and maximum temperature simulation over complex terrain. *Ann. Appl. Stat.* **7** 588–612. MR3086432

KLEIBER, W. and NYCHKA, D. (2012). Nonstationary modeling for multivariate spatial processes. *J. Multivariate Anal.* **112** 76–91. MR2957287

KLEIBER, W. and PORCU, E. (2015). Nonstationary matrix covariances: Compact support, long range dependence and quasi-arithmetic constructions. *Stoch. Environ. Res. Risk Assess.* **29** 193–204.

KÜNSCH, H. R., PAPRITZ, A. and BASSI, F. (1997). Generalized cross-covariances and their estimation. *Math. Geol.* **29** 779–799. MR1470669

LARK, R. M. (2003). Two robust estimators of the cross-variogram for multivariate geostatistical analysis of soil properties. *European J. Soil Sci.* **54** 187–201.

LI, B., GENTON, M. G. and SHERMAN, M. (2008). Testing the covariance structure of multivariate random fields. *Biometrika* **95** 813–829. MR2461213

LI, B. and ZHANG, H. (2011). An approach to modeling asymmetric multivariate spatial covariance structures. *J. Multivariate Anal.* **102** 1445–1453. MR2819961

LIM, C. Y. and STEIN, M. (2008). Properties of spatial cross-periodograms using fixed-domain asymptotics. *J. Multivariate Anal.* **99** 1962–1984. MR2466546

LONG, A. E. and MYERS, D. E. (1997). A new form of the cokriging equations. *Math. Geol.* **29** 685–703. MR1460688

MA, C. (2011a). Vector random fields with second-order moments or second-order increments. *Stoch. Anal. Appl.* **29** 197–215. MR2774237

MA, C. (2011b). Covariance matrices for second-order vector random fields in space and time. *IEEE Trans. Signal Process.* **59** 2160–2168. MR2816489

MA, C. (2011c). Vector random fields with long-range dependence. *Fractals* **19** 249–258. MR2803242

MA, C. (2011d). A class of variogram matrices for vector random fields in space and/or time. *Math. Geosci.* **43** 229–242. MR2764536

MA, C. (2012). Stationary and isotropic vector random fields on spheres. *Math. Geosci.* **44** 765–778. MR2956272

MA, C. (2013a). Student's *t* vector random fields with power-law and log-law decaying direct and cross covariances. *Stoch. Anal. Appl.* **31** 167–182. MR3007889

MA, C. (2013b). Mittag–Leffler vector random fields with Mittag–Leffler direct and cross covariance functions. *Ann. Inst. Statist. Math.* **65** 941–958. MR3105803

MAJUMDAR, A. and GELFAND, A. E. (2007). Multivariate spatial modeling for geostatistical data using convolved covariance functions. *Math. Geol.* **39** 225–245. MR2324633

MAJUMDAR, A., PAUL, D. and BAUTISTA, D. (2010). A generalized convolution model for multivariate nonstationary spatial processes. *Statist. Sinica* **20** 675–695. MR2682636

MARDIA, K. V. and GOODALL, C. R. (1993). Spatial–temporal analysis of multivariate environmental monitoring data. In *Multivariate Environmental Statistics*. *North-Holland Ser. Statist. Probab.* **6** 347–386. North-Holland, Amsterdam. MR1268443

MEARNS, L. O., GUTOWSKI, W. J., JONES, R., LEUNG, A. M., MCGINNIS, B., NUNES, Y. and QIAN, Y. (2009). A regional climate change assessment program for North America. *Eos, Transactions, American Geophysical Union* **90** 311–312.

MYERS, D. E. (1982). Matrix formulation of co-kriging. *J. Internat. Assoc. Math. Geol.* **14** 249–257. MR0660192

MYERS, D. E. (1983). Estimation of linear combinations and co-kriging. *J. Internat. Assoc. Math. Geol.* **15** 633–637. MR0719602

MYERS, D. E. (1991). Pseudo-cross variograms, positive-definiteness, and cokriging. *Math. Geol.* **23** 805–816. MR1121438

MYERS, D. E. (1992). Kriging, co-kriging, radial basis functions and the role of positive definiteness. *Comput. Math. Appl.* **24** 139–148. MR1190311

NARCOWICH, F. J. and WARD, J. D. (1994). Generalized Hermite interpolation via matrix-valued conditionally positive definite functions. *Math. Comp.* **63** 661–687. MR1254147

NORTH, G. R., WANG, J. and GENTON, M. G. (2011). Correlation models for temperature fields. *J. Climate* **24** 5850–5862.

OEHLERT, G. W. (1993). Regional trends in sulfate wet deposition. *J. Amer. Statist. Assoc.* **88** 390–399.

PACIOREK, C. J. and SCHERVISH, M. J. (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics* **17** 483–506. MR2240939

PAPRITZ, A., KÜNSCH, H. R. and WEBSTER, R. (1993). On the pseudo cross-variogram. *Math. Geol.* **25** 1015–1026. MR1246957

PASCUAL, F. and ZHANG, H. (2006). Estimation of linear correlation coefficient of two correlated spatial processes. *Sankhyā* **68** 307–325. MR2303086

PETERSON, T. C. and VOSE, R. S. (1997). An overview of the Global Historical Climatology Network temperature database. *Bull. Amer. Meteorol. Soc.* **78** 2837–2849.

PORCU, E., BEVILACQUA, M. and GENTON, M. G. (2014). Spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere. Unpublished manuscript.

PORCU, E., GREGORI, P. and MATEU, J. (2006). Nonseparable stationary anisotropic space–time covariance functions. *Stoch. Environ. Res. Risk Assess.* **21** 113–122. MR2307631

PORCU, E. and ZASTAVNYI, V. (2011). Characterization theorems for some classes of covariance functions associated to vector valued random fields. *J. Multivariate Anal.* **102** 1293–1301. MR2811618

PORCU, E., DALEY, D. J., BUHMANN, M. and BEVILACQUA, M. (2013). Radial basis functions with compact support for multivariate geostatistics. *Stoch. Environ. Res. Risk Assess.* **27** 909–922.

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

ROUHANI, S. and WACKERNAGEL, H. (1990). Multivariate geostatistical approach to space–time data analysis. *Water Resour. Res.* **26** 585–591.

ROYLE, J. A. and BERLINER, L. M. (1999). A hierarchical approach to multivariate spatial modeling and prediction. *J. Agric. Biol. Environ. Stat.* **4** 29–56. MR1812239

RUIZ-MEDINA, M. D. and PORCU, E. (2015). Equivalence of Gaussian measures of multivariate random fields. *Stoch. Environ. Res. Risk Assess.* **29** 325–334.

SAIN, S. R. and CRESSIE, N. (2007). A spatial model for multivariate lattice data. *J. Econometrics* **140** 226–259. MR2395923

SAMPSON, P. D. and GUTTORP, P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *J. Amer. Statist. Assoc.* **87** 108–119.

SANG, H., JUN, M. and HUANG, J. Z. (2011). Covariance approximation for large multivariate spatial data sets with an application to multiple climate model errors. *Ann. Appl. Stat.* **5** 2519–2548. MR2907125

SCHEUERER, M. and SCHLATHER, M. (2012). Covariance models for divergence-free and curl-free random vector fields. *Stoch. Models* **28** 433–451. MR2959449

SCHLATHER, M. (2010). Some covariance models based on normal scale mixtures. *Bernoulli* **16** 780–797. MR2730648

SCHMIDT, A. M. and GELFAND, A. E. (2003). A Bayesian coregionalization approach for multivariate pollutant data. *J. Geophys. Res.* **108** 1–9.

SHMARYAN, L. E. and JOURNEL, A. G. (1999). Two Markov models and their application. *Math. Geol.* **31** 965–988. MR1717034

STEIN, M. L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. *Springer Series in Statistics*. Springer, New York. MR1697409

STEIN, M. L. (2005). Nonstationary spatial covariance functions. Technical Report 21, Univ. Chicago, CISES.

SUBRAMANYAM, A. and PANDALAI, H. S. (2008). Data configurations and the cokriging system: Simplification by screen effects. *Math. Geosci.* **40** 425–443. MR2401515

SUN, Y., LI, B. and GENTON, M. G. (2012). Geostatistics for large datasets. In *Space–Time Processes and Challenges Related to Environmental Problems* **207** (E. Porcu, J. M. Montero and M. Schlather, eds.) 55–77. Springer, Berlin.

VARGAS-GUZMÁN, J. A., WARRICK, A. W. and MYERS, D. E. (1999). Multivariate correlation in the framework of support and spatial scales of variability. *Math. Geol.* **31** 85–104.

VARGAS-GUZMÁN, J. A., WARRICK, A. W. and MYERS, D. E. (2002). Coregionalization by linear combination of nonorthogonal components. *Math. Geol.* **34** 405–419. MR1951789

VER HOEF, J. M. and BARRY, R. P. (1998). Constructing and fitting models for cokriging and multivariable spatial prediction. *J. Statist. Plann. Inference* **69** 275–294. MR1631328

VER HOEF, J. M. and CRESSIE, N. (1993). Multivariable spatial prediction. *Math. Geol.* **25** 219–240. MR1206187

VER HOEF, J. M., CRESSIE, N. and BARRY, R. P. (2004). Flexible spatial models for kriging and cokriging using moving averages and the fast Fourier transform (FFT). *J. Comput. Graph. Statist.* **13** 265–282. MR2063985

WACKERNAGEL, H. (1994). Cokriging versus kriging in regionalized multivariate data analysis. *Geoderma* **62** 83–92.

WACKERNAGEL, H. (2003). *Multivariate Geostatistics*: *An Introduction with Applications*, 3rd ed. Springer, Berlin.

ZHANG, H. (2007). Maximum-likelihood estimation for multivariate spatial linear coregionalization models. *Environmetrics* **18** 125–139. MR2345650