Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: https://www.tandfonline.com/loi/uasa20

A Valid Matérn Class of Cross-Covariance Functions for Multivariate Random Fields With Any Number of Components

Tatiyana V. Apanasovich , Marc G. Genton & Ying Sun

To cite this article: Tatiyana V. Apanasovich , Marc G. Genton & Ying Sun (2012) A Valid Matérn Class of Cross-Covariance Functions for Multivariate Random Fields With Any Number of Components, Journal of the American Statistical Association, 107:497, 180-193, DOI: 10.1080/01621459.2011.643197

To link to this article: https://doi.org/10.1080/01621459.2011.643197

Published online: 11 Jun 2012.

Submit your article to this journal

Article views: 593

View related articles

Citing articles: 37 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=uasa20

A Valid Matérn Class of Cross-Covariance Functions for Multivariate Random Fields With Any Number of Components

Tatiyana V. APANASOVICH, Marc G. GENTON, and Ying SUN

We introduce a valid parametric family of cross-covariance functions for multivariate spatial random fields where each component has a covariance function from a well-celebrated Matérn class. Unlike previous attempts, our model indeed allows for various smoothnesses and rates of correlation decay for any number of vector components. We present the conditions on the parameter space that result in valid models with varying degrees of complexity. We discuss practical implementations, including reparameterizations to reflect the conditions on the parameter space and an iterative algorithm to increase the computational efficiency. We perform various Monte Carlo simulation experiments to explore the performances of our approach in terms of estimation and cokriging. The application of the proposed multivariate Matérn model is illustrated on two meteorological datasets: temperature/pressure over the Pacific Northwest (bivariate) and wind/temperature/pressure in Oklahoma (trivariate). In the latter case, our flexible trivariate Matérn model is valid and yields better predictive scores compared with a parsimonious model with common scale parameters.

KEY WORDS: Cokriging; Correlation decay; Multivariate; Smoothness; Spatial; Valid cross-covariance.

1. INTRODUCTION

Multivariate data indexed by spatial coordinates have become ubiquitous in a large number of applications, for instance, in environmental and climate sciences, to name but a few. This has prompted a renewed interest in multivariate random fields in recent years for modeling purpose. Specifically, consider a p-dimensional multivariate random field $\mathbf{Z}(\mathbf{s}) = \{Z_1(\mathbf{s}), \ldots, Z_p(\mathbf{s})\}^{\mathrm{T}}$ defined on a spatial region $\mathcal{D} \subset \mathbb{R}^d$, $d \geq 1$, where $Z_i(\mathbf{s})$ represents the ith variable, $i = 1, \ldots, p$, at location $\mathbf{s} \in \mathcal{D}$. If $\mathbf{Z}$ is assumed to be Gaussian, then only its mean and cross-covariance functions need to be described to fully specify $\mathbf{Z}$. Here, we assume that $\mathbf{Z}$ has mean vector zero and is second-order stationary (or just stationary); hence, the cross-covariance functions

$$\operatorname{cov}\{Z_i(\mathbf{s}_1), Z_j(\mathbf{s}_2)\} = C_{ij}(\mathbf{s}_1 - \mathbf{s}_2), \quad i, j = 1, \ldots, p, \quad \mathbf{s}_1, \mathbf{s}_2 \in \mathcal{D}, \tag{1}$$

depend only on the separation vector $\mathbf{s}_1 - \mathbf{s}_2$; see, for example, Wackernagel (2003) and references therein. Stationarity can be thought of as an invariance property under the translation of coordinates. Isotropic fields are stationary fields invariant under rotations and reflections so that $C_{ij}(\mathbf{h}_1) = C_{ij}(\mathbf{h}_2)$ whenever $\|\mathbf{h}_1\| = \|\mathbf{h}_2\|$, where $\|\cdot\|$ denotes the Euclidean norm. Isotropy or even stationarity are not always realistic, but sometimes are satisfactory working assumptions. Moreover, these models serve as building blocks for more sophisticated anisotropic and nonstationary covariances; for instance, see Mateu, Porcu, and Gregori (2008), Porcu, Mateu, and Bevilacqua (2007), and Nott and Dunsmuir (2002). Our goal is to construct valid, and flexible, parametric cross-covariance functions for (1). This is a challenging task because of the well-known requirement that the covariance matrix $\boldsymbol{\Sigma}$ of the random vector $(\mathbf{Z}(\mathbf{s}_1)^{\mathrm{T}}, \ldots, \mathbf{Z}(\mathbf{s}_n)^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{np}$ should be nonnegative definite; that is, $\mathbf{a}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{a} \geq 0$ for any nonzero vector $\mathbf{a} \in \mathbb{R}^{np}$, any spatial locations $\mathbf{s}_1, \ldots, \mathbf{s}_n$, and any integer n.

Various approaches to construct valid cross-covariance models can be found in the literature. The simplest model is given by separable (also called intrinsic) cross-covariance functions (Mardia and Goodall 1993) for which $C_{ij}(\mathbf{s}_1 - \mathbf{s}_2) = \rho(\mathbf{s}_1 - \mathbf{s}_2)A_{ij}$, where $\mathbf{A} = \{A_{ij}\}_{i,j=1}^{p}$ is a $p \times p$ nonnegative definite matrix and $\rho(\cdot)$ is a valid stationary correlation function. The main limitation of the separable model is that the covariance between components at each spatial location has the same shape independently of the relative displacement of the locations. A popular approach is to use a linear model of coregionalization (LMC) for stationary random fields (Goulard and Votz 1992; Schmidt and Gelfand 2003; Wackernagel 2003; Gelfand et al. 2004; Zhang 2007). It consists in representing the multivariate random field as a linear combination of r independent univariate random fields. The resulting cross-covariance functions take the form $C_{ij}(\mathbf{s}_1 - \mathbf{s}_2) = \sum_{k=1}^{r} \rho_k(\mathbf{s}_1 - \mathbf{s}_2)A_{ik}A_{jk}$ for an integer $1 \leq r \leq p$, where $\rho_k(\cdot)$ are valid stationary correlation functions and $\mathbf{A} = \{A_{ij}\}_{i,j=1}^{p}$ is a $p \times r$ full rank matrix. The main drawback of the LMC is that the smoothness of any component of the multivariate random field is that of the roughest underlying univariate random field. Other approaches to valid cross-covariance models include kernel convolution (Ver Hoef and Barry 1998) and covariance convolution (Gaspari and Cohn 1999; Majumdar and Gelfand 2007) methods, both of which generally require numerical integration. Recently, Apanasovich and Genton (2010) proposed an approach based on latent dimensions for the construction of valid cross-covariance

Tatiyana V. Apanasovich is Assistant Professor, Division of Biostatistics, Thomas Jefferson University, Philadelphia, PA 19107 (E-mail: *tatiyana.apanasovich@jefferson.edu*). Marc G. Genton is Professor, Department of Statistics, Texas A&M University, College Station, TX 77843-3143 (E-mail: *genton@stat.tamu.edu*). Ying Sun is postdoctoral associate, Statistical and Applied Mathematical Sciences Institute, 19 T.W. Alexander Drive, Research Triangle Park, NC 27709-4006 (E-mail: *sunwards@samsi.info*). This publication is based in part on work supported by award no. KUS-C1-016-04 made by King Abdullah University of Science and Technology (KAUST) and by NSF grants DMS-1007504 and DMS-0707106. The authors thank the editor, two referees, and Tilmann Gneiting for their helpful comments and suggestions.

© 2012 American Statistical Association
Journal of the American Statistical Association
March 2012, Vol. 107, No. 497, Theory and Methods
DOI: 10.1080/01621459.2011.643197

functions. Their method produces flexible models that can control various types of nonseparability and can also allow for components with different smoothnesses, although the implementation of the latter is not straightforward.

The quest for valid cross-covariance models that can allow for different smoothnesses of the components of the multivariate random field is therefore important and has practical applications. In the univariate setting, that is, when $p = 1$, there are many available covariance models (Gneiting, Genton, and Guttorp 2007) but the Matérn family (Matérn 1960) has found widespread interest in recent years, perhaps due to its strong support from Stein (1999); see also Guttorp and Gneiting (2006) for a historical account of this model. It defines the spatial covariance function of the univariate random field at two locations $\mathbf{s}_1$ and $\mathbf{s}_2$ as $\sigma^2 M(\mathbf{s}_1 - \mathbf{s}_2|\nu, \alpha)$, where σ^2 is the marginal variance and

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

where K_ν is a modified Bessel function of the second kind, $\nu > 0$ is a smoothness parameter, and $\alpha > 0$ is a scale parameter. Here, α measures how quickly the correlation of the random field decays with distance, with larger α corresponding to a faster decay (keeping ν fixed). The ratio $1/\alpha$ is sometimes referred to as a correlation length and is closely related to what is known as the practical range defined as the distance at which the correlations are nearly 0, say 0.05. The parameter ν measures the smoothness of the random field with larger values of ν corresponding to smoother fields. In particular, the random field is m times mean square differentiable if and only if $\nu > m$. Very recently, Gneiting, Kleiber, and Schlather (2010) have introduced a class of Matérn cross-covariance functions, a multivariate version of (2) defined by

$$C_{ij}(\mathbf{h}) = \sigma_{ij} M(\mathbf{h}|\nu_{ij}, \alpha_{ij}), \quad i, j = 1, \ldots, p, \quad \mathbf{h} \in \mathbb{R}^d, \quad (3)$$

with co-located covariance coefficients σ_{ij}, smoothness parameters ν_{ij}, and scale parameters α_{ij}. Although these authors provided a full characterization of the values of σ_{ij}, ν_{ij}, and α_{ij} that result in a valid bivariate Matérn model, that is, when $p = 2$, they were only able to give conditions for $p \geq 3$ in the particular case of common scale parameters and restricted smoothness parameters ($\nu_{ij} = (\nu_{ii} + \nu_{jj})/2$), the so-called parsimonious model. The main goal of this article is to give a general characterization of the parameters that yield a valid multivariate Matérn model for any number of components.

The remainder of the article is organized as follows. In Section 2, we discuss conditions on the multivariate Matérn model that lead to valid cross-covariance functions (3). In particular, we consider what we call a flexible multivariate Matérn model for any number of components as well as various special cases. We also study the bivariate Matérn model in detail. Practical implementations, including reparameterizations to reflect the conditions on the parameter space and an iterative algorithm to increase the computational efficiency, are discussed in Section 3. We perform various Monte Carlo simulation experiments to explore the performances of our approach in Section 4. We illustrate in Section 5 the application of the proposed flexible multivariate Matérn model on two meteorological datasets: temperature/pressure over the Pacific Northwest (bivariate) and wind/temperature/pressure in Oklahoma (trivariate). In the latter case, our flexible trivariate Matérn model is valid and yields better predictive scores compared with a parsimonious model with common scale parameters. The proofs of the theoretical results are collected in the Appendix.

2. MAIN THEORETICAL RESULTS

We discuss conditions on the multivariate Matérn model that lead to valid cross-covariance functions (3). Even though our flexible model (by opposition to the parsimonious model by Gneiting et al. 2010) is not full in the sense of a criterion, we show in our numerical studies that it is flexible enough to perform well in practical applications.

2.1 Flexible Multivariate Matérn Model

First we consider a multivariate Matérn model (3), which is the least parsimonious model we offer, where all parameters are allowed to vary with i and j.

Theorem 1. The flexible multivariate Matérn model provides a valid structure if there exists $\Delta_A \geq 0$ such that

(i) $\nu_{ij} - (\nu_{ii} + \nu_{jj})/2 = \Delta_A(1 - A_{ij})$, $\quad i, j = 1, \ldots, p$, where $0 \leq A_{ij} \leq 1$ form a valid correlation matrix;
(ii) $-\alpha_{ij}^2$, $i, j = 1, \ldots, p$, form a conditional nonnegative definite matrix; and
(iii) $\sigma_{ij}\alpha_{ij}^{2\Delta_A + \nu_{ii} + \nu_{jj}} \Gamma(\nu_{ij} + d/2)/[\Gamma\{(\nu_{ii} + \nu_{jj})/2 + d/2\}\Gamma(\nu_{ij})]$, $\quad i, j = 1, \ldots, p$, form a nonnegative definite matrix.

The proof of Theorem 1 is deferred to the Appendix, where we also provide some background on nonnegative definite, conditional nonnegative definite, and infinitely divisible matrices. The following remark gives some examples of structures of $\{\alpha_{ij}\}_{i,j=1}^p$ that part (ii) permits.

Remark 1. Some examples of a collection of $\{\alpha_{ij}\}_{i,j=1}^p$ that imply the condition (ii) of Theorem 1 are

(a) $\alpha_{ij}^2 = (\alpha_{ii}^2 + \alpha_{jj}^2)/2 + \tau(\alpha_{ii} - \alpha_{jj})^2$, $0 \leq \tau < \infty$;
(b) $\alpha_{ij} = \max(\alpha_{ii}, \alpha_{jj})$;
(c) $\alpha_{ij}^2 = (\alpha_{ii}^2 + \alpha_{jj}^2)/2 + \Delta_B(1 - B_{ij})$, where $\Delta_B \geq 0$, and $0 \leq B_{ij} \leq 1$ form a valid correlation matrix.

We sketch the proofs in the Appendix.

Next, we make several remarks about the constraints that the conditions of Theorem 1 impose on the parameters.

Remark 2. Condition (i) implies that the smoothness parameter in the cross-covariance of two processes is greater or equal to the average of the smoothness parameters in the covariance functions of the corresponding marginal processes, that is, $\nu_{ij} \geq (\nu_{ii} + \nu_{jj})/2$. The constraint is not surprising since it is necessary in the bivariate case that $\nu_{12} \geq (\nu_{11} + \nu_{22})/2$, as shown in Gneiting et al. (2010).

Remark 3. Condition (ii) implies that the squared scale parameter in the cross-covariance of two processes is greater or equal to the average of the squared scale parameters of the corresponding marginal processes, that is, $\alpha_{ij}^2 \geq (\alpha_{ii}^2 + \alpha_{jj}^2)/2$.

Corollary 1 considers the special case of $\alpha_{ij}^2 < (\alpha_{ii}^2 + \alpha_{jj}^2)/2$ when $\nu_{ij} = (\nu_{ii} + \nu_{jj})/2$ (see Remark 5).

In all applications we considered, it is reasonable to assume that the cross-correlation decays faster than the correlation of at least one of the corresponding marginal processes. Therefore, $\alpha_{ij}^2 \geq (\alpha_{ii}^2 + \alpha_{jj}^2)/2$ agrees well with our applications. However, Corollary 1, part (b), provides a starting point for models where one would need that $\alpha_{ij}^2 < (\alpha_{ii}^2 + \alpha_{jj}^2)/2$.

Remark 4. Condition (iii) implies the following constraints on the cross-correlations:

$$\sigma_{ij}^2/(\sigma_{ii}\sigma_{jj}) \leq \prod_{i=1}^{3} \tau_{ij}^{(i)} \leq 1, \quad i, j = 1, \ldots, p, \tag{4}$$

$$\tau_{ij}^{(1)} = \frac{\mathcal{B}(\nu_{ij}, d/2)^2}{\mathcal{B}\{(\nu_{ii} + \nu_{jj})/2, d/2\}^2}, \quad \tau_{ij}^{(2)} = \left(\frac{\alpha_{ii}\alpha_{jj}}{\alpha_{ij}^2}\right)^{2\Delta_A},$$

$$\tau_{ij}^{(3)} = \frac{\Gamma^2\{(\nu_{ii} + \nu_{jj})/2\}/\alpha_{ij}^{2(\nu_{ii}+\nu_{jj})}}{\{\Gamma(\nu_{ii})/\alpha_{ii}^{2\nu_{ii}}\}\{\Gamma(\nu_{jj})/\alpha_{ij}^{2\nu_{jj}}\}},$$

where $\mathcal{B}(\cdot, \cdot)$ is the Beta function. The proof is given in the Appendix.

The constraints on cross-correlations (see Remark 4) seem to be unavoidable, as Gneiting et al. (2010) showed in their criterion for the bivariate Matérn model. Basically, the upper bounds (4) on the cross-correlations depend on how much the smoothness and scale parameters deviate from the corresponding parameters of the two relevant marginal processes. The upper bound is 1 when the parameters are the same for all correlations and cross-correlations, as in a case of separability. We discuss the upper bounds further in Section 2.3.

Theorem 1 provides guidance on the parameterization of the cross-covariances. One just needs to choose the parameterization for nonnegative definite matrices mentioned in (i)–(iii). We discuss the parameterization in detail in Section 3.1.

2.2 Special Cases of Flexible Multivariate Matérn Model

Zhang (2004) showed that the parameters σ^2 and α, assuming ν is known, of a univariate Matérn covariance function cannot be estimated consistently under infill asymptotics. Thus, constraining scale or smoothness parameters is not really restrictive in practical applications. Next, we discuss cases when either ν_{ij} or α_{ij}^2, $i \neq j$, are closely related to parameters from the marginal correlation functions of their counterparts.

Similar to the bivariate criterion, the case $\nu_{ij} = (\nu_{ii} + \nu_{jj})/2$ deserves a special attention.

Corollary 1. The multivariate Matérn model with $\nu_{ij} = (\nu_{ii} + \nu_{jj})/2$ is valid if either

(a) $-\alpha_{ij}^2$ form a conditional nonnegative definite matrix and $\sigma_{ij}\alpha_{ij}^{2\nu_{ij}}/\Gamma(\nu_{ij})$ nonnegative definite matrices; or
(b) $-\alpha_{ij}^{-2}$ form a conditional nonnegative definite matrix and $\sigma_{ij}/\{\alpha_{ij}^d\Gamma(\nu_{ij})\}$ nonnegative definite matrices;

$i, j = 1, \ldots, p$.

Part (a) follows directly from Theorem 1 by letting $\Delta_A = 0$. The proof of part (b) is deferred to the Appendix.

Remark 5. The constraint on α_{ij} from part (b) of Corollary 1 implies that the squared cross-correlation length is greater or equal to the average of the squared correlation lengths of the corresponding processes, that is, $\alpha_{ij}^{-2} \geq (\alpha_{ii}^{-2} + \alpha_{jj}^{-2})/2$. Moreover, the constrained sets of α_{ij} from (a) and (b) overlap in the parameter space only when they are all equal, $\alpha_{ij} = \alpha$ (see Remark 8).

Remark 6. Part (b) implies the following constraints on the cross-correlations:

$$\frac{\sigma_{ij}^2}{\sigma_{ii}\sigma_{jj}} \leq \tau_{ij}^{(4)} \leq 1, \quad \tau_{ij}^{(4)} = \frac{\Gamma(\nu_{ij})^2\alpha_{ij}^{2d}}{\Gamma(\nu_{ii})\alpha_{ii}^d\Gamma(\nu_{jj})\alpha_{jj}^d}, \quad i, j = 1, \ldots, p. \tag{5}$$

The following corollary lists sufficient conditions on parameters to provide valid cross-covariances in cases where the vector components share either the same smoothness or the same scale parameters.

Corollary 2. The following multivariate Matérn models are valid:

(a) $\alpha_{ij} = \alpha > 0$, $-\nu_{ij}$ form conditionally nonnegative matrices and $\sigma_{ij}\Gamma(\nu_{ij} + d/2)/\Gamma(\nu_{ij})$ nonnegative definite matrices; or
(b) $\nu_{ij} = \nu > 0$, if either:
- $-\alpha_{ij}^2$ form a conditionally nonnegative definite matrix and $\sigma_{ij}\alpha_{ij}^{2\nu}$ form nonnegative definite matrices; or
- $-\alpha_{ij}^{-2}$ form a conditionally nonnegative definite matrix and $\sigma_{ij}/\alpha_{ij}^d$ form nonnegative definite matrices;

$i, j = 1, \ldots, p$.

Part (b) follows directly from Corollary 1. The proof of part (a) is deferred to the Appendix.

Remark 7. The special case of Corollary 2, part (a), $\nu_{ij} = (\nu_{ii} + \nu_{jj})/2, \alpha_{ij} = \alpha_{jj} = \alpha$, for any $i, j = 1, \ldots, p$, is consistent with Theorem 1 from Gneiting et al. (2010) and is called "parsimonious" by these authors.

The following corollary provides sufficient conditions for the validity of the multivariate Matérn model when the parameters for cross-correlations are averages of parameters from the corresponding vector components.

Corollary 3. The multivariate Matérn model is valid for the following cases:

(a) $\nu_{ij} = (\nu_{ii} + \nu_{jj})/2$, $\alpha_{ij}^2 = (\alpha_{ii}^2 + \alpha_{jj}^2)/2$, and $\sigma_{ij}\alpha_{ij}^{2\nu_{ij}}/\Gamma(\nu_{ij})$ form nonnegative definite matrices;
(b) $\nu_{ij} = (\nu_{ii} + \nu_{jj})/2$, $\alpha_{ij}^{-2} = (\alpha_{ii}^{-2} + \alpha_{jj}^{-2})/2$, and $\sigma_{ij}/\{\Gamma(\nu_{ij})\alpha_{ij}^d\}$ form nonnegative definite matrices;

$i, j = 1, \ldots, p$.

The claims follow directly from Corollary 1 and Remark 1, part (a).

2.3 The Bivariate Matérn Model

The bivariate case is the simplest of all multivariate Matérn models. In fact, the conditions for the validity of the structure reduce to one single inequality discussed by Gneiting et al. (2010). Here, we compare sufficient conditions from Theorem 1 and Corollaries 1 and 2 with the optimal ones provided by Gneiting et al. (2010) for the case of two vector components.

Corollary 4. The bivariate Matérn model $\sigma_{ij}M(\mathbf{h}|\nu_{ij},\alpha_{ij})$, $i,j=1,2$, provides a valid structure if:

(a) $\nu_{12}>(\nu_{11}+\nu_{22})/2$, $\alpha_{12}^2\geq(\alpha_{11}^2+\alpha_{22}^2)/2$, and $\sigma_{12}^2/(\sigma_{11}\sigma_{22})\leq\prod_{k=1}^{3}\tau_{12}^{(k)}$, where $\Delta_A=\nu_{12}-(\nu_{11}+\nu_{22})/2$; or
(b) $\nu_{12}=(\nu_{11}+\nu_{22})/2$, $\alpha_{12}^2\geq(\alpha_{11}^2+\alpha_{22}^2)/2$, and $\sigma_{12}^2/(\sigma_{11}\sigma_{22})\leq\tau_{12}^{(3)}$; or
(c) $\nu_{12}=(\nu_{11}+\nu_{22})/2$, $\alpha_{12}^{-2}\geq(\alpha_{11}^{-2}+\alpha_{22}^{-2})/2$, and $\sigma_{12}^2/(\sigma_{11}\sigma_{22})\leq\tau_{12}^{(4)}$; or
(d) $\nu_{12}\geq(\nu_{11}+\nu_{22})/2$, $\alpha_{12}=\alpha_{11}=\alpha_{22}$, and $\sigma_{12}^2/(\sigma_{11}\sigma_{22})\leq\frac{\Gamma(\nu_{12})^2\Gamma(\nu_{11}+d/2)\Gamma(\nu_{22}+d/2)}{\Gamma(\nu_{12}+d/2)^2\Gamma(\nu_{11})\Gamma(\nu_{22})}$;

where $\tau_{12}^{(k)}$, $k=1,2,3$, and $\tau_{12}^{(4)}$ are defined in (4) and (5), respectively.

Remark 8. If $\alpha_{12}^2\geq(\alpha_{11}^2+\alpha_{22}^2)/2$, then $\alpha_{12}^{-2}<(\alpha_{11}^{-2}+\alpha_{22}^{-2})/2$, unless $\alpha_{ij}=\alpha$, $i,j=1,2$.

We compare the optimal upper bound on $\rho_{12}=|\sigma_{12}/(\sigma_{11}\sigma_{22})^{1/2}|$ from Gneiting et al. (2010) with the one resulting from our sufficient conditions, guided by the parameters from the marginal models in our bivariate numerical studies: $\alpha_{11}=0.020$, $\alpha_{22}=0.010$, $\alpha_{12}=0.016$, $\nu_{11}=1.6$, $\nu_{22}=0.6$, and $\nu_{12}=1.3$. Each valid combination of smoothing and scale parameters has to constrain the cross-correlation parameter, ρ_{12}, as proved in Gneiting et al. (2010). Our results provide sufficient conditions; hence, it is expected that the upper bound for ρ_{12} from our theory is less or equal to the optimal upper bound from Gneiting et al. (2010). Therefore, a significant difference between the upper bounds indicates a somewhat less flexible construction that our theory offers. However, as our numerical studies show, that lack of flexibly does not seem to affect the predictive performance of the model. All six pictures in Figure 1 contain two curves depicting our upper bound (dashed with circles) and the optimal bound

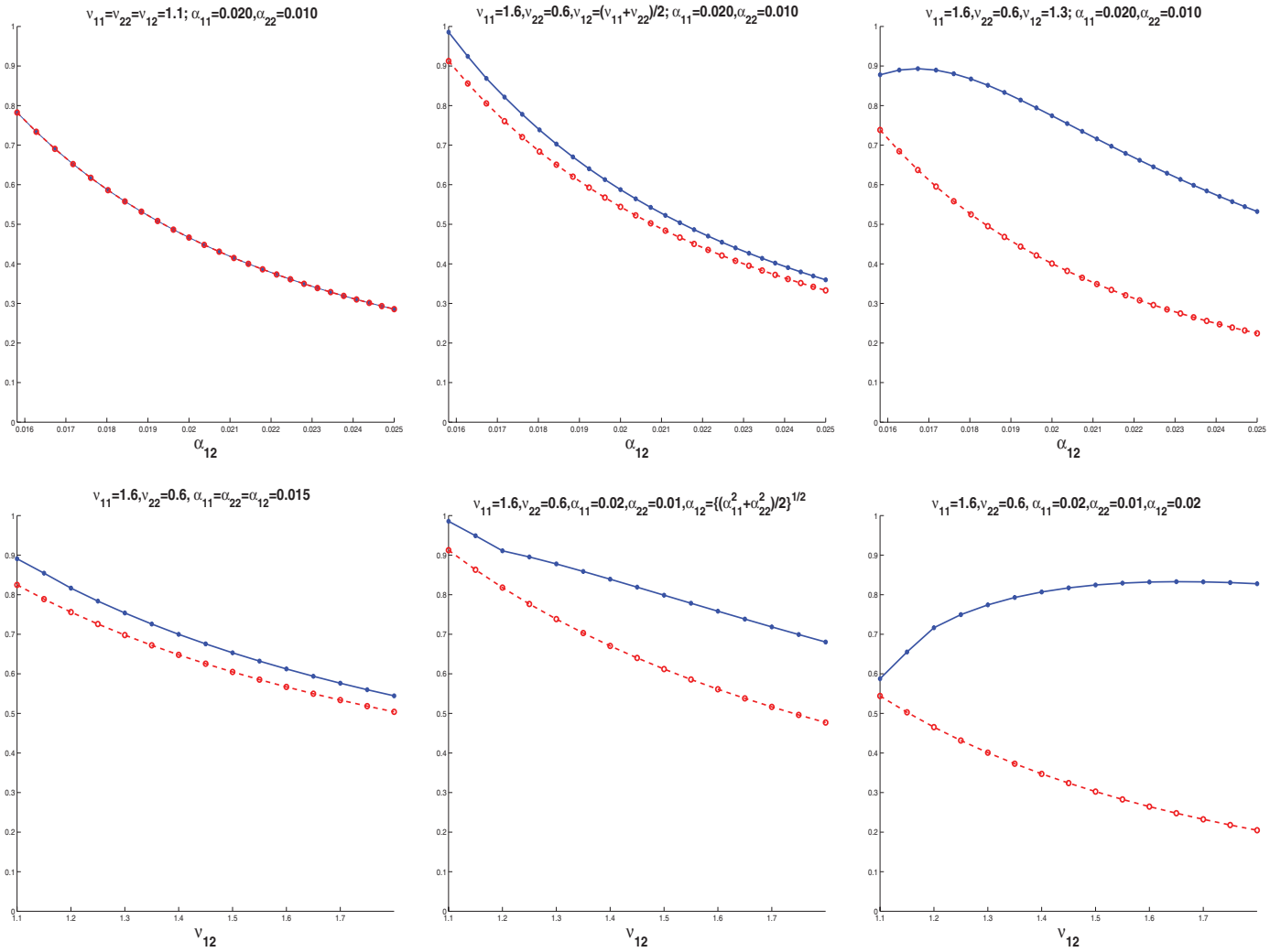

Figure 1. Upper bounds on Matérn cross-correlations for a bivariate random field: upper bound that agrees with the proposed model (dashed with circles) and optimal upper bound (solid with dots) for various combinations of parameters. (The online version of this figure is in color.)

(solid with dots). The upper panel of Figure 1 contains three plots, depicting upper bounds as a function of α_{12}, for three cases reflecting relationships between ν_{ij}'s $(i, j = 1, 2)$. The first, second, and third plots (from left to right) portrait the cases $\nu_{12} = \nu_{11} = \nu_{22} = 1.1$, $\nu_{12} = (\nu_{11} + \nu_{22})/2 = 1.1$, and $\nu_{12} = 1.3 > (\nu_{11} + \nu_{22})/2$, respectively. One can observe that the upper bounds perfectly agree when $\nu_{12} = \nu_{11} = \nu_{22}$, agree well in the case $\nu_{12} = (\nu_{11} + \nu_{22})/2$, and disagree when ν_{12} differs significantly from $(\nu_{11} + \nu_{22})/2$, and the size of the disagreement is not affected by α_{12}^2 unless it is far from $(\alpha_{11}^2 + \alpha_{22}^2)/2$. The bottom panel of Figure 1 contains three plots, depicting upper bounds as a function of ν_{12}, for three cases reflecting relationships between α_{ij}^2. Similar to the upper panel, the first, second, and third plots (from left to right) portrait the cases $\alpha_{ij} = 0.015$, $\alpha_{12} = \{(\alpha_{11}^2 + \alpha_{22}^2)/2\}^{1/2}$, and the initial $\alpha_{12}^2 > (\alpha_{11}^2 + \alpha_{22}^2)/2$. One can observe that the upper bounds agree well in the case $\alpha_{12}^2 = \alpha_{11}^2 = \alpha_{22}^2$ and disagree when α_{ij}^2 are different, and the size of the disagreement is affected by both: the distance of α_{12}^2 from $(\alpha_{11}^2 + \alpha_{22}^2)/2$ and of ν_{12} from $(\nu_{11} + \nu_{22})/2$.

Remark 9. It is difficult to estimate how much flexibility is lost in the multivariate case with sufficient (but not necessary) conditions our theorems provide by comparing only pairwise upper bounds. This is because, in general, the validity of all pairwise correlations does not guarantee the validity of the entire structure. However, if the optimal model and ours share the marginal smoothing and scale parameters, it is evident that our estimates of smoothness and scale for the cross-covariances (ν_{12} and α_{12}^2) will tend to be closer to the mean of their marginal counterparts compared with the optimal ones, to prevent ρ_{12} from being grossly underestimated (see Section 5.1 for illustration).

3. PRACTICAL IMPLEMENTATION

In this section, we discuss how to use the theoretical developments from Section 2 in practice. Specifically, we describe the parameterization that agrees with the main theorem and propose an iterative estimation scheme for improved computational stability.

3.1 Parameterization

To generate a parameterization that agrees with Theorem 1, one needs to specify three matrices and a parameter $\Delta_A \geq 0$, in addition to the parameters from the marginal covariances. Next, we describe the general path to a valid parameterization.

Parameterization of α_{ij}. We choose parameterization (c) from Remark 1. Let $\mathbf{R}_B$ be a valid correlation matrix with nonnegative entries. Then,

$$\alpha_{ij}^2 = (\alpha_{ii}^2 + \alpha_{jj}^2)/2 + \Delta_B(1 - R_{B,ij}), \quad \Delta_B \geq 0, \quad i, j = 1, \ldots, p. \tag{6}$$

When $\Delta_B \equiv 0$, $\mathbf{R}_B$ becomes nonidentifiable, so we distinguish two cases: $\Delta_B \equiv 0$ and no $\mathbf{R}_B$, and $\Delta_B > 0$ and $\mathbf{R}_B$ is a valid correlation matrix with nonnegative entries. When $\mathbf{R}_B$ is equicorrelated, it becomes nonidentifiable as well; see Remark 10. When $R_{B,ij} = 0$, $i \neq j$, $\alpha_{ij}^2 - (\alpha_{ii}^2 + \alpha_{jj}^2)/2 = \Delta_B$, and when $R_{B,ij} = 1, i \neq j, \alpha_{ij}^2 - (\alpha_{ii}^2 + \alpha_{jj}^2)/2 = 0$. Therefore, $R_{B,ij}$ can be interpreted as parameters that add flexibility to model deviation of $\alpha_{ij}^2, i \neq j$, from $(\alpha_{ii}^2 + \alpha_{jj}^2)/2$. The parameter Δ_B can be interpreted as a parameter responsible for the range of deviations of α_{ij}^2 from $(\alpha_{ii}^2 + \alpha_{jj}^2)/2$, where $\Delta_B = 0$ means no deviation at all.

Parameterization of ν_{ij}. Similarly, for ν_{ij}, to be consistent with (i) in Theorem 1, we introduce a valid correlation matrix with nonnegative entries, $\mathbf{R}_A$, a parameter $\Delta_A > 0$ and let

$$\nu_{ij} = \frac{\nu_{ii} + \nu_{jj}}{2} + \Delta_A(1 - R_{A,ij}) \geq 0, \quad i, j = 1, \ldots, p. \tag{7}$$

Parameterization of σ_{ij}. We introduce $\mathbf{R}_V$, a valid correlation matrix, $W_{ii} > 0, i = 1, \ldots, p$, and let

$$\sigma_{ij} = W_{ii} W_{jj} R_{V,ij} \alpha_{ij}^{-2\Delta_A - (\nu_{ii} + \nu_{jj})} \Gamma\{(\nu_{ii} + \nu_{jj})/2 + d/2\} \times \Gamma(\nu_{ij}) / \Gamma(\nu_{ij} + d/2), \quad i, j = 1, \ldots, p. \tag{8}$$

Other parameterization schemes follow directly from the corollaries and can be combined with the parameterizations described above.

One can choose any parameterization for the correlation matrices $\mathbf{R}_L, L \in \{A, B, V\}$. Those matrices can be modeled using the techniques proposed in Apanasovich and Genton (2010). For example, $R_{L,ij} = \exp(-||\xi_{L,i} - \xi_{L,j}||)$, for vectors $\xi_{L,i} \in \mathbb{R}^k$, $1 \leq k \leq p$, under constraints discussed in Apanasovich and Genton (2010).

Remark 10. In the case of a small number of variables, p, one can use equicorrelated $\mathbf{R}_L$s so that $R_{L,ij} = \rho_L, i \neq j, L \in \{A, B, V\}$. Redefine $\Delta_L := \Delta_L(1 - \rho_L)$, for $L \in \{A, B\}$, so that $\alpha_{ij}^2 = (\alpha_{ii}^2 + \alpha_{jj}^2)/2 + \Delta_B$ and $\nu_{ij} = (\nu_{ii} + \nu_{jj})/2 + \Delta_A, i \neq j$, $i, j = 1, \ldots, p$.

The least flexible parameterization (because it requires no extra parameters to model smoothness and scale parameters from cross-correlations), when $\Delta_A = \Delta_B = 0$, follows from Corollary 3. For example, for part (a),

$$\nu_{ij} = \frac{\nu_{ii} + \nu_{jj}}{2}, \quad \alpha_{ij}^2 = \frac{\alpha_{ii}^2 + \alpha_{jj}^2}{2},$$
$$\sigma_{ij} = (\sigma_{ii}\sigma_{jj})^{1/2} \frac{\alpha_{ii}^{\nu_{ii}} \alpha_{jj}^{\nu_{jj}}}{\alpha_{ij}^{2\nu_{ij}}} \frac{\Gamma(\nu_{ij})}{\Gamma^{1/2}(\nu_{ii})\Gamma^{1/2}(\nu_{jj})} R_{V,ij}. \tag{9}$$

One can introduce the constraints involving σ_{ij} directly into the covariance structure. For example, the following cross-covariance function is valid and follows from Corollary 3, part (a):

$$C_{ij}(\mathbf{h}) = \frac{(\sigma_{ii}\sigma_{jj})^{1/2} R_{V,ij} \Gamma\{(\nu_{ii} + \nu_{jj})/2\} / \{\Gamma(\nu_{ii})\Gamma(\nu_{jj})\}^{1/2}}{\{(\alpha_{ii}^2 + \alpha_{jj}^2)/2\}^{(\nu_{ii} + \nu_{jj})/2} / \left\{\alpha_{ii}^{2\nu_{ii}} \alpha_{jj}^{2\nu_{jj}}\right\}^{1/2}} \times M\left\{\mathbf{h} \middle| (\nu_{ii} + \nu_{jj})/2, \sqrt{(\alpha_{ii}^2 + \alpha_{jj}^2)/2}\right\},$$

for $i, j = 1, \ldots, p$, where $R_{V,ij}$ are cross-correlations and σ_{ii} are variances in the traditional sense.

3.2 Estimation

In this section, we discuss the algorithms we successfully implemented in our numerical studies. The maximum likelihood and restricted maximum likelihood methods are generally considered the best options for estimating model parameters.

Table 1. Summary statistics of parameter estimation of the bivariate ($p = 2$) Matérn model over 1000 replicates

Parameter	True	Min.	Q_1	Median	Mean	Q_3	Max.
α_{11}	0.02	0.0095	0.0193	0.0217	0.0225	0.0248	0.0511
α_{22}	0.01	0.0041	0.0099	0.0124	0.0133	0.0153	0.1139
α_{12}	0.0158	0.0074	0.0165	0.0186	0.0194	0.0215	0.0826
ν_{11}	1.6	1.1580	1.5670	1.7060	1.7600	1.8730	4.2040
ν_{22}	0.6	0.4082	0.6018	0.6760	0.7508	0.7737	2.8400
ν_{12}	1.3	0.9209	1.2220	1.3590	1.5330	1.6490	4.3800
σ_{11}	1	0.6035	0.9381	1.0520	1.0530	1.1600	1.6190
σ_{22}	1	0.5942	0.9575	1.0500	1.0510	1.1380	1.6350
σ_{12}	−0.497	−0.9921	−0.5352	−0.4425	−0.4352	−0.3403	0.0000

However, those methods can be impractical for large datasets, and likelihood-based approximations to the objective function exist, for example, composite likelihood (Lindsay 1988; Apanasovich et al. 2008). Our models can be fit using any optimization method one finds reasonable when balancing computational complexity and statistical efficiency.

Since the main goal of this article is to introduce a novel parameterization, rather than advance computational machinery, we use a computationally efficient algorithm that we describe next. Our estimates can be used as initial guesses in more computationally intense algorithms, for example, numerical maximization of the likelihood, using the limited-memory quasi-Newton bound-constrained optimization method of Byrd et al. (1995).

Here is our algorithm:

- *Step 1.* Fit the marginal models and estimate α_{ii}, ν_{ii}, and σ_{ii}, $i = 1, \ldots, p$.
- *Step 2.* Choose parametric forms for $\mathbf{R}_A$, $\mathbf{R}_B$, and $\mathbf{R}_V$, parameterized by vectors $\rho_{A,ij}$, $\rho_{B,ij}$, and $\rho_{V,ij}$, respectively. Express α_{ij}^2 through α_{ii}^2, $R_{B,ij}$, and Δ_B using (6), and ν_{ij} through ν_{ii}, $R_{A,ij}$, and Δ_A using (7). Let σ_{ij} be as in (8) by using $R_{V,ij}$ and parameters W_{ii}, $i = 1, \ldots, p$.
- *Step 3.* Fit the joint model to estimate Δ_A, Δ_B, and $\rho_{L,ij}$, $L \in \{A, B, V\}$, while keeping α_{ii}, ν_{ii}, and σ_{ii} fixed from Step 1.

Remark 11. In the case of bivariate data, Step 2 can be replaced by the following:

- *Step* 2^*. Define $\nu_{12} = (\nu_{11} + \nu_{22})/2 + \Delta_A$ and $\alpha_{12}^2 = (\alpha_{11}^2 + \alpha_{22}^2)/2 + \Delta_B$ for parameters $\Delta_A, \Delta_B \geq 0$. Let $\sigma_{12}^2 = \rho_V^2(\sigma_{11}\sigma_{22})\prod_{k=1}^3 \tau_{12}^{(k)}$, for a parameter $\rho_V^2 \leq 1$, where $\tau_{12}^{(k)}$, $k = 1, 2, 3$ are defined by (4).

Remark 12. From our numerical studies, we noted that the problem of practical nonidentifiability can occur, especially for low-dimension p. In this case, we recommend to introduce a penalty term into the optimization procedure to make the solution unique, say, $-(\Delta_A^2 + \Delta_B^2)$ in the maximization. Then, out of all possible combinations of $(\Delta_A, \Delta_B, \mathbf{R}_A, \mathbf{R}_B)$, the solution will be the one for which $(\Delta_A^2 + \Delta_B^2)$ is the smallest.

4. MONTE CARLO SIMULATIONS

To explore the performance of the proposed model, we conduct simulation studies for the cases of two variables ($p = 2$) and three variables ($p = 3$). The simulation scenarios are motivated by a meteorological dataset discussed by Gneiting et al. (2010). It consists of temperature and pressure observations, as well as forecasts, at 157 locations in the North American Pacific Northwest. In our simulation studies, we use these same 157 locations and generate a bivariate or trivariate spatial Gaussian random field with multivariate Matérn cross-covariance structure parameterized in Section 3.1. For $p = 2$, as stated in Remark 11, only one parameter ρ_V is needed to parameterize the correlation matrix $\mathbf{R}_V$, and the choice of parameter values in the simulations is consistent with our estimates for the dataset in Section 5.1; for example, the correlation parameter $\rho_{12} = \sigma_{12}/\sqrt{\sigma_{11}\sigma_{22}}$ is chosen to be -0.497, representing the negative correlation between temperatures and pressures. Following our advice (Remark 10), for $p = 3$, we choose the correlation matrices to be equicorrelated and parameterized by Δ_L for $L \in \{A, B, V\}$, respectively. The choice of the parameters is also motivated by a real dataset, the wind/temperature/pressure application in Section 5.2 with two negative correlations and one positive correlation, indicating how the three variables are correlated. Then, we apply the algorithm outlined in Section 3.2 to the maximum likelihood estimation (MLE). However, the second step in the algorithm for the bivariate case was modified as advised in Remark 11, and in the case of three-vector components, we followed Remark 12. The empirical distribution of each parameter estimate is summarized in Tables 1 and 2 based on 1000 replicates.

As can be observed from Tables 1 and 2, we allow the scale parameters α_{ij} in the simulations to differ for both $p = 2$ and $p = 3$ cases. For all parameters, the medians and means of 1000 estimates are close to their true values, that is, they have small biases. Also, the small spreads show the low variability of the estimates.

A simulation study for $p = 2$ on one special case where $\nu_{12} = (\nu_{11} + \nu_{22})/2$ and $\alpha_{12}^2 = (\alpha_{11}^2 + \alpha_{22}^2)/2$ is also conducted. The parameterization of σ_{12} follows Equation (9). The simulation results for 1000 replicates are reported in Table 3 and also show a good performance of the parameter estimation.

To further study the performance of our model in terms of prediction, we compare the mean squared prediction error (MSPE) and the mean absolute error (MAE) of our flexible Matérn model with those of the particular model above and the parsimonious model (see Remark 7) for $p = 2$ by simulation. Here, we use the same true parameter values as in Table 1, but randomly select 57 out of 157 locations for prediction purpose. We use the

Table 2. Summary statistics of parameter estimation of the trivariate ($p = 3$) Matérn model over 1000 replicates

Parameter	True	Min.	Q_1	Median	Mean	Q_3	Max.
α_{11}	0.01	0.0035	0.0095	0.0111	0.0115	0.0130	0.0237
α_{22}	0.02	0.0076	0.0188	0.0236	0.0273	0.0300	0.2659
α_{33}	0.03	0.0013	0.0288	0.0326	0.0691	0.0722	0.8220
α_{12}	0.0205	0.0093	0.0185	0.0222	0.0252	0.0269	0.1943
α_{13}	0.0263	0.0071	0.0254	0.0363	0.0626	0.0614	0.8220
α_{23}	0.0282	0.0138	0.0303	0.0404	0.0668	0.0648	0.8220
ν_{11}	1.2	0.9953	1.1800	1.2560	1.2770	1.3480	1.9090
ν_{22}	0.6	0.3552	0.5915	0.6938	0.9028	0.8264	7.3200
ν_{33}	0.3	0.0515	0.3235	0.3677	1.2800	0.8636	7.3100
ν_{12}	1.093	0.8709	1.0690	1.1300	1.1790	1.2080	5.2270
ν_{13}	1.092	0.8449	1.0850	1.1490	1.3090	1.2510	6.3270
ν_{23}	0.990	0.5943	0.8659	1.1740	1.2450	1.3390	5.2270
σ_{11}	1	0.3354	0.7560	0.9158	0.9460	1.0990	3.2600
σ_{22}	1	0.5932	0.8726	0.9770	0.9812	1.0680	1.4650
σ_{33}	1	0.2341	0.8059	0.9265	0.9126	1.0290	1.4900
σ_{12}	−0.286	−0.7036	−0.3306	−0.2565	−0.2583	−0.1916	0.0000
σ_{13}	−0.181	−0.4858	−0.2692	−0.1715	−0.1711	−0.0518	0.0000
σ_{23}	0.274	0.0000	0.1103	0.2313	0.2916	0.3669	0.5199

three aforementioned estimated Matérn cross-covariance models to predict the primary variable at the 57 left-out locations by cokriging techniques; for a recent account, see Furrer and Genton (2011) and references therein. With 1000 replications, the boxplots of the MSPE and MAE for the three models are shown in the top panels of Figure 2. Our model has a smaller average MSPE, 0.300, than 0.329 and 0.819, the average MSPE from the particular and the parsimonious models, respectively, as well as a smaller average MAE. The boxplots also clearly support the better performance of our flexible model in terms of MSPE and MAE. Since these two measures only depend on the correlation function but do not evaluate the fitted variance, we compare some predictive scores to evaluate the predictive performance by taking the prediction variance into account. We consider two predictive scores, the logarithmic score (LogS) and the continuous ranked probability score (CRPS), proposed by Gneiting and Raftery (2007) and Gneiting, Balabdaoui, and Raftery (2007) and described by Zhang and Wang (2010) in the context of large datasets. The boxplots of LogS and CRPS from the simulations are shown in the bottom panels of Figure 2. Both medians are the lowest for our model, which indicates that our model has lower prediction error and hence better predictive performance when evaluated by LogS and CRPS.

5. APPLICATIONS

5.1 Bivariate Temperature/Pressure Spatial Field

For the case of two variables, we illustrate the use of the bivariate Matérn model on the same dataset used by Gneiting et al. (2010) and show that we obtain a better fit. The data consist of temperature and pressure observations, as well as forecasts at 157 locations in the North American Pacific Northwest. Similarly to these authors, we consider the bivariate spatial error random field that arises as forecast minus observation. We aim at fitting a random field model for pressure and temperature errors, allowing them to have different scale parameters. For this dataset, temperature and pressure errors are strongly negatively correlated and the co-located empirical correlation coefficient is −0.47.

We consider a Gaussian random field $\mathbf{Y}(\mathbf{s}) = (Y_1(\mathbf{s}), Y_2(\mathbf{s}))^{\mathrm{T}}$, where $\mathbf{s} \in \mathbb{R}^2$, Y_1 denotes temperature, and Y_2 denotes pressure. We fit the bivariate Matérn model (3) with nugget effects τ_1^2 and τ_2^2 to account for measurement errors to the centered

Table 3. Summary statistics of parameter estimation of a particular bivariate ($p = 2$) Matérn model over 1000 replicates

Parameter	True	Min.	Q_1	Median	Mean	Q_3	Max.
α_{11}	0.02	0.0108	0.0188	0.0222	0.0230	0.0261	0.0684
α_{22}	0.01	0.0034	0.0098	0.0121	0.0129	0.0152	0.0485
α_{12}	0.0158	0.0092	0.0155	0.0181	0.0188	0.0213	0.0498
ν_{11}	1.2	0.8311	1.1620	1.2860	1.3460	1.4490	7.6190
ν_{22}	0.6	0.3870	0.5937	0.6600	0.6979	0.7577	3.5830
ν_{12}	0.900	0.6408	0.8923	0.9854	1.0220	1.1040	4.1830
σ_{11}	1	0.4878	0.8448	0.9629	0.9738	1.0830	1.6720
σ_{22}	1	0.4728	0.8059	0.9373	0.9479	1.0770	1.6540
σ_{12}	−0.828	−1.3690	−0.8807	−0.7841	−0.7931	−0.6833	−0.3352

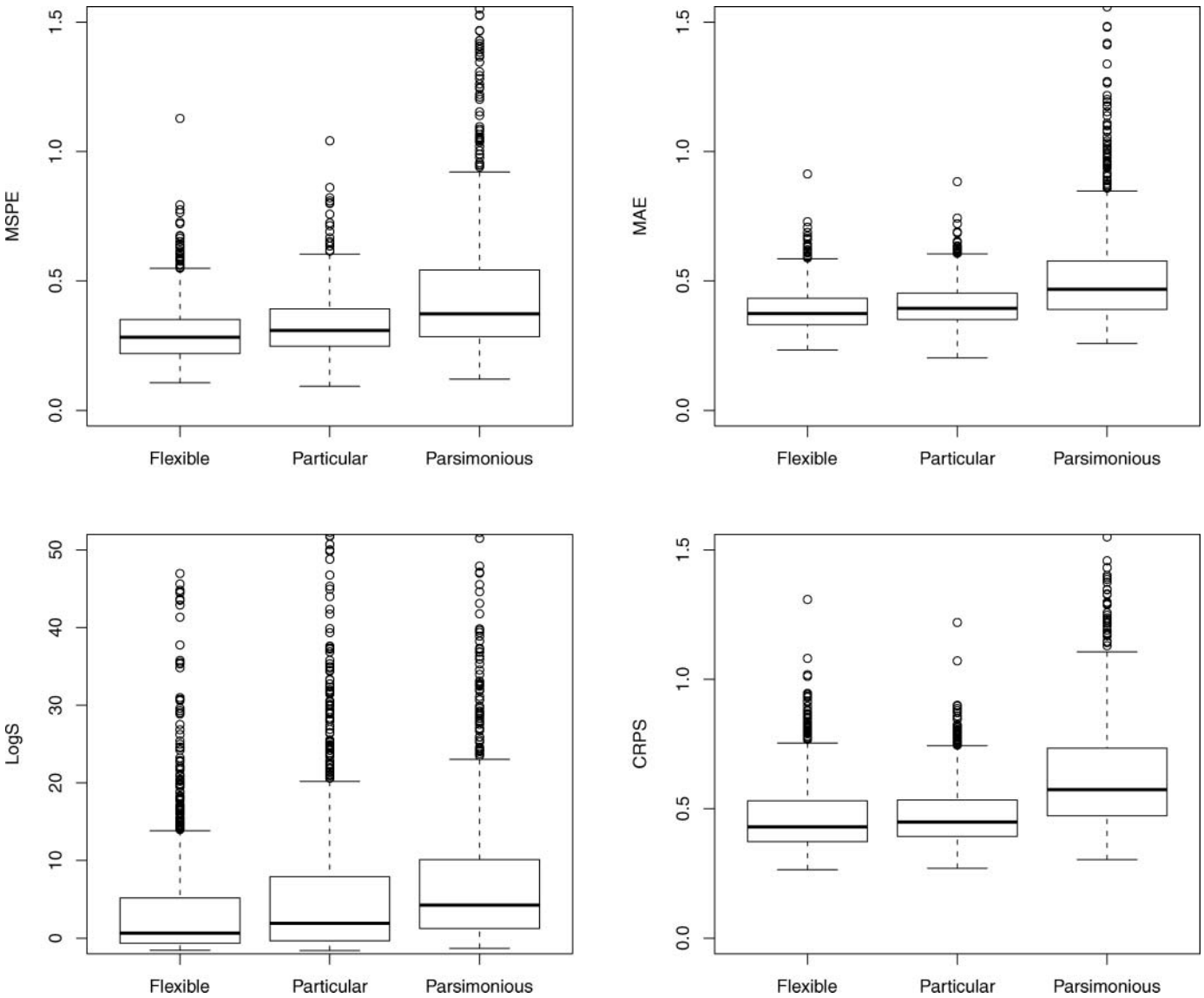

Figure 2. Boxplots of MSPE, MAE, LogS, and CRPS for the flexible, particular, and the parsimonious bivariate Matérn models with 1000 replicates.

field, $\mathbf{Y}(\mathbf{s}) - \bar{\mathbf{Y}}$, where $\bar{\mathbf{Y}}$ is the vector of sample means. Our bivariate covariance model thus becomes

$$C_{11}(\mathbf{h}) = \sigma_{11} M(\mathbf{h}|\nu_{11}, \alpha_{11}) + \tau_1^2 \mathcal{I}(\mathbf{h} = \mathbf{0}),$$
$$C_{22}(\mathbf{h}) = \sigma_{22} M(\mathbf{h}|\nu_{22}, \alpha_{22}) + \tau_2^2 \mathcal{I}(\mathbf{h} = \mathbf{0}),$$
$$C_{12}(\mathbf{h}) = C_{21}(\mathbf{h}) = \sigma_{12} M(\mathbf{h}|\nu_{12}, \alpha_{12}),$$

where $\mathcal{I}$ denotes the indicator function.

In the estimation procedure, we apply the proposed algorithm to the standardized observations and then transform the obtained estimates back to the original scale. For $p = 2$, to estimate the Matérn cross-covariance model, only one parameter ρ_V is needed to parameterize the correlation matrix $\mathbf{R}_V$, as described by the algorithm outlined in Section 3.2. Numerical optimization of the flexible Matérn likelihood leads to the estimates in Table 4. The marginal correlation and cross-correlation fits are depicted in Figure 3. As can be observed from the parameter estimates in Table 4 or the fits in Figure 3, the error field is rough for temperature, $\hat{\nu}_{11} = 0.59$, and smooth for pressure, $\hat{\nu}_{22} = 1.61$, which agrees well with the laws of physics. For instance, North, Wang, and Genton (2011) have shown that $\nu_{11} = 1$ is implied by energy-balance climate models for temperature fields and have argued that rougher fields, that is, $\nu_{11} < 1$, could be expected due to vegetation and terrains. Note that in Gneiting et al.'s (2010) fit, the means of Ys were not estimated but set to zeroes instead, which, in theory, leads to larger variance estimators and has an influence on scale and smoothness parameter estimators as well. Therefore, to better compare our flexible model with the full and the parsimonious models of Gneiting et al. (2010), we refit their models by plugging in the sample means as the MLEs in the optimization procedures. The parameter estimates and the log-likelihood of the full and the parsimonious models are also shown in Table 4, and the marginal correlation and cross-correlation fits are depicted in Figure 3. For the flexible and the full models, we estimate the marginal models first, and conditional on the marginal parameters, we estimate the remaining cross-covariance parameters. They are thus based on the same estimated marginal models

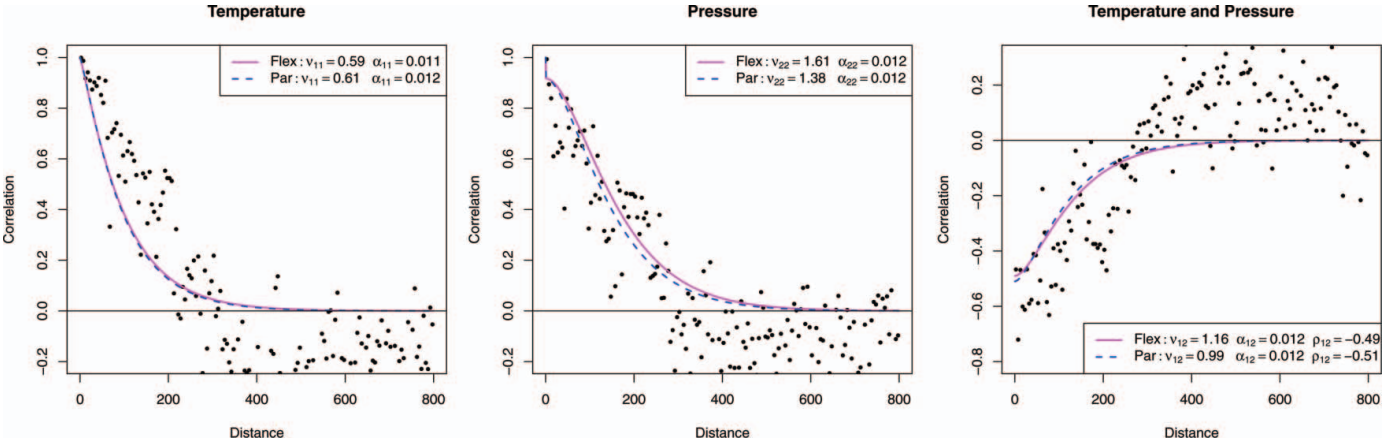

Figure 3. Marginal correlation and cross-correlation fits of temperature and pressure data: solid curves for the flexible bivariate Matérn model and dashed curves for the parsimonious bivariate Matérn model. (The online version of this figure is in color.)

and are only different at the joint model estimation stage. Table 4 shows that the values of the log-likelihood for our flexible model and Gneiting et al.'s (2010) full model are close, while the parsimonious model has a slightly lower value of the log-likelihood due to the common scale constraint. As expected (see Remark 9), our $\hat{\nu}_{12} = 1.16$ is closer to $(\hat{\nu}_{11} + \hat{\nu}_{22})/2 = 1.10$ than the corresponding parameter from the full model, 1.50. Similarly, $1/\hat{\alpha}_{12} = 81.3$ is closer to $\sqrt{2/(\hat{\alpha}_{11}^2 + \hat{\alpha}_{22}^2)} = 86.6$ than the estimate from the full model, 70.9. However, as one can observe from the log-likelihoods, the differences in parameter estimates do not affect the fit significantly. Now, for the full and the parsimonious models, by estimating the mean vector, the log-likelihood is further increased compared with the values reported by Gneiting et al. (2010). Here, the co-located cross-correlation is defined as $\rho_{12} = \sigma_{12}/\sqrt{\sigma_{11}\sigma_{22}}$, and its estimate in our fit is strongly negative, $\hat{\rho}_{12} = -0.49$. Thus, unlike the zero mean model, we appear to be capturing the features of the joint spatial distribution of the two variables better (with a larger log-likelihood) with the flexible bivariate Matérn model, rather than with a common scale model, and with nonzero means. The fits in Figure 3 are also visually better.

5.2 Trivariate Wind/Temperature/Pressure Spatial Field

We now illustrate the use of the proposed flexible multivariate Matérn model for both model fitting and prediction on a meteorological dataset that consists of wind speed, temperature, and pressure observations at 120 locations in Oklahoma, as depicted in Figure 4. Accurate forecasts of wind speed are of crucial importance in many applications, such as harvesting electricity from wind energy (Genton and Hering 2007; Hering and Genton 2010) and severe weather warnings for the general public. For this dataset, we randomly select 100 locations for model fitting and leave out the other 20 locations to evaluate the wind speed prediction performance. The observations are measured at 2:20 p.m. local time, on March 8, 2010.

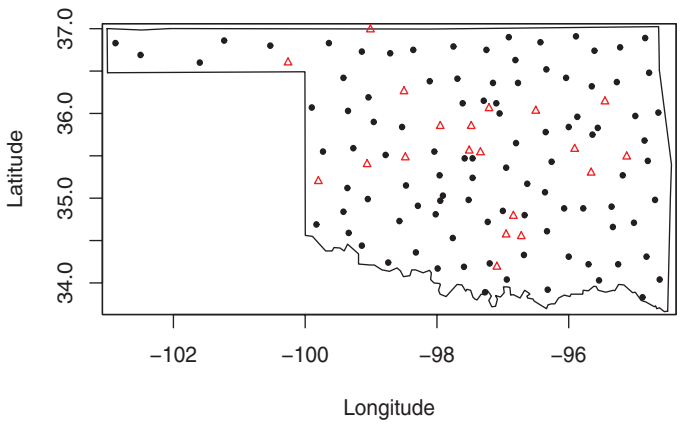

Figure 4. The locations of 120 weather stations in Oklahoma. Black dots (100 locations) are the sites for model fitting and red triangles (20 locations) are the complementary sites for prediction. (The online version of this figure is in color.)

For the three variables at the 100 locations, we consider a mean zero Gaussian random field $\mathbf{Y}(\mathbf{s}) = (Y_1(\mathbf{s}), Y_2(\mathbf{s}), Y_3(\mathbf{s}))^{\mathrm{T}}$, where $\mathbf{s} \in \mathbb{R}^2$ and Y_i, $i = 1, 2, 3$, denotes wind speed, temperature, and pressure, respectively. We first remove a quadratic trend of longitude, latitude, and elevation for each variable and test the multivariate normality of the residuals. The Shapiro–Wilks test gives a p-value equal to 0.3. With such a

Table 4. Estimates of parameters for the bivariate Matérn model (Flex: our fit; Full: Gneiting et al.'s (2010) fit with nonzero means) and the parsimonious Matérn model (Pars) with nonzero means applied to the temperature and pressure data with corresponding values of the log-likelihood (Loglik). Units are degree Celsius for temperature, Pascal for pressure, and kilometer for distances

Model	$\hat{\sigma}_{11}$	$\hat{\sigma}_{22}$	$\hat{\rho}_{12}$	$\hat{\nu}_{11}$	$\hat{\nu}_{22}$	$\hat{\nu}_{12}$	$1/\hat{\alpha}_{11}$	$1/\hat{\alpha}_{22}$	$1/\hat{\alpha}_{12}$	$\hat{\tau}_1$	$\hat{\tau}_2$	Loglik
Flex	6.81	51099	−0.49	0.59	1.61	1.16	93.2	81.3	81.3	0	68.0	−1263.4
Full	6.81	51099	−0.54	0.59	1.61	1.50	93.2	81.3	70.9	0	68.0	−1263.3
Pars	6.81	51099	−0.51	0.61	1.38	0.99	86.7	86.7	86.7	0	68.0	−1263.8

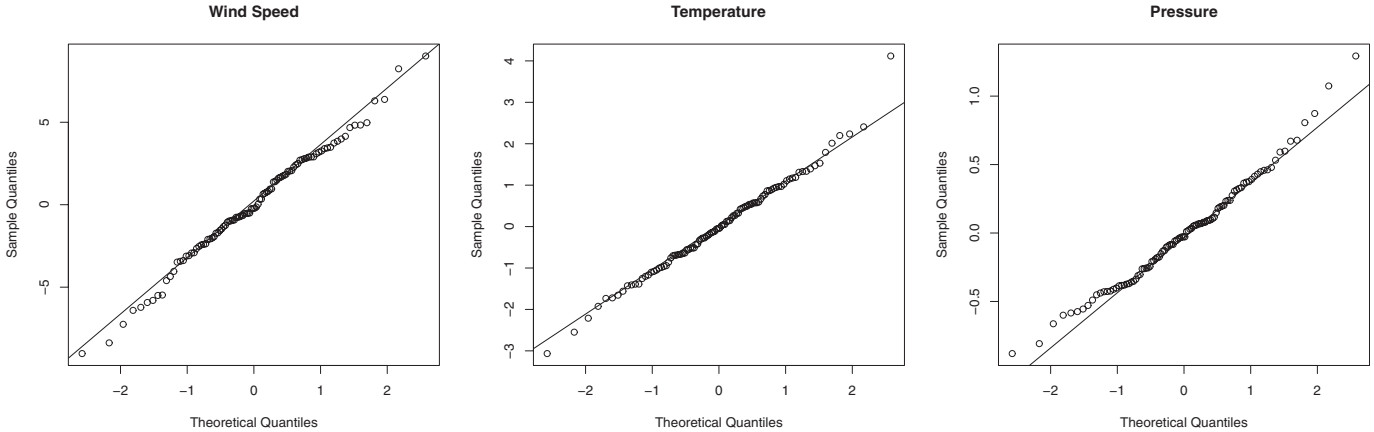

Figure 5. Marginal Q-Q plots of wind speed, temperature, and pressure.

large p-value, we fail to reject the multivariate normality. The Q-Q plots in Figure 5 also show marginal normality.

Here, we aim at fitting a random field model for wind speed, temperature, and pressure errors after removing a trend. For this dataset, the wind speed errors are positively correlated with temperature errors, but negatively correlated with pressure errors. Temperature and pressure errors are negatively correlated. The co-located empirical correlation coefficients are 0.36 between wind speed and temperature, -0.24 between wind speed and pressure, and -0.25 between temperature and pressure.

To estimate the Matérn cross-covariance model for $p = 3$, we choose the correlation matrices $\mathbf{R}_A$ and $\mathbf{R}_B$ to be equicorrelated and parameterized by ρ_L for $L \in \{A, B\}$, respectively, and let the correlation matrix $\mathbf{R}_V$ be parameterized by $\rho_{V,12}$, $\rho_{V,13}$, and $\rho_{V,23}$. Then, we apply the algorithm outlined in Section 3.2 to the MLE. Numerical optimization of the flexible Matérn likelihood leads to the estimates in Table 5. The smoothness parameter is estimated as $\hat{\nu}_{11} = 0.77$ for wind speed, which, as expected, is rougher than that for temperature, $\hat{\nu}_{22} = 1.32$, and pressure, $\hat{\nu}_{33} = 1.97$. The estimates of the co-located correlation

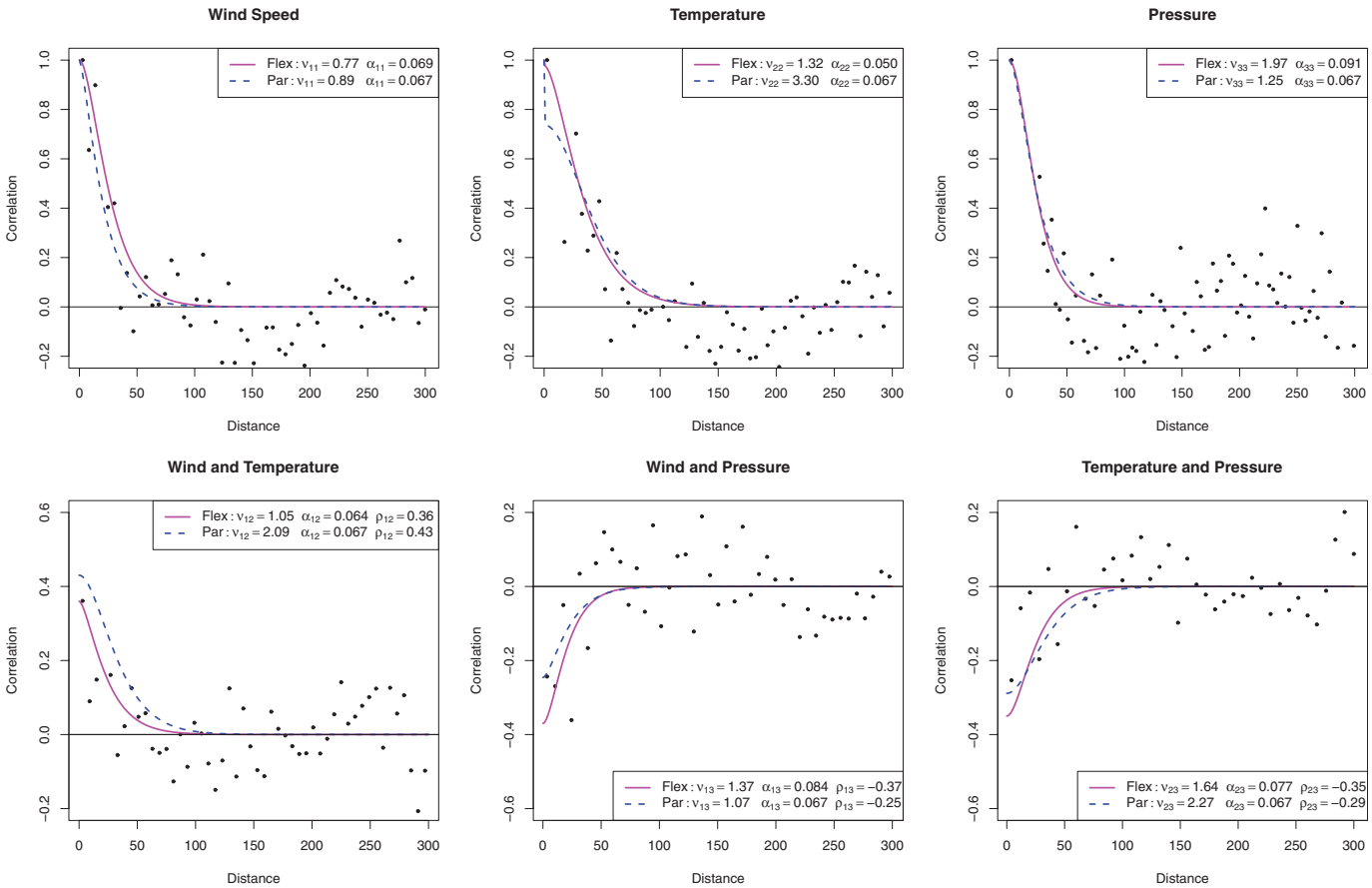

Figure 6. Marginal correlation and cross-correlation fits for wind speed, temperature, and pressure: solid curves for the flexible trivariate Matérn model and dashed curves for the parsimonious trivariate Matérn model. (The online version of this figure is in color.)

Table 5. Maximum likelihood estimates of parameters for our flexible trivariate Matérn model applied to the wind speed, temperature, and pressure data. Units are miles per hour for wind speed, degree Celsius for temperature, Pascal for pressure, and kilometer for distances

$\hat{\sigma}_{11}$	$\hat{\sigma}_{22}$	$\hat{\sigma}_{33}$	$\hat{\rho}_{12}$	$\hat{\rho}_{13}$	$\hat{\rho}_{23}$	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{\tau}_3$			
11.65	1.40	0.16	0.36	−0.35	−0.37	0	0.18	0.01			
$\hat{\nu}_{11}$	$\hat{\nu}_{22}$	$\hat{\nu}_{33}$	$\hat{\nu}_{12}$	$\hat{\nu}_{13}$	$\hat{\nu}_{23}$	$1/\hat{\alpha}_{11}$	$1/\hat{\alpha}_{22}$	$1/\hat{\alpha}_{33}$	$1/\hat{\alpha}_{12}$	$1/\hat{\alpha}_{13}$	$1/\hat{\alpha}_{23}$
0.77	1.32	1.97	1.05	1.37	1.64	14.5	20.0	11.0	15.6	11.9	13.0

coefficients are in agreement with the empirical ones. Thus, we capture the main features of the joint spatial distribution of the three variables. The marginal correlation and cross-correlation fits for wind speed, temperature, and pressure are depicted in Figure 6 with solid curves.

5.3 Cokriging Predictions for $p = 3$

We now use the estimated flexible trivariate Matérn model to perform the prediction at the 20 locations that were not included in model fitting. The primary variable of interest is wind speed. Suppose we observe the temperature and pressure at all the 120 locations but only observe the wind speed at the 100 locations. Then, we may use the estimated Matérn cross-covariance model to predict the wind speed at the 20 locations by cokriging techniques. Since we know the real observed wind speed, we can then compute the predictive scores, MSPE, MAE, LogS, and CRPS, described in Section 4, and compare them with those obtained by using the model with common scale parameters.

The parsimonious multivariate Matérn model in Gneiting et al. (2010) assumes $\nu_{ij} = (\nu_i + \nu_j)/2, \alpha_{11} = \cdots = \alpha_{pp} = \alpha$, and $\alpha_{ij} = \alpha$ for $1 \le i \ne j \le p$. Then, it provides a valid structure if the correlation matrix consists of

$$\rho_{ij} = \beta_{ij} \frac{\nu_{ii}^{1/2} \nu_{jj}^{1/2}}{(\nu_{ii} + \nu_{jj})/2}, \quad i, j = 1, \ldots, p,$$

where the matrix $\{\beta_{ij}\}_{i,j=1}^{p}$ with diagonal elements $\beta_{ii} = 1$ for $i = 1, \ldots, p$ and off-diagonal β_{ij} is symmetric and nonnegative definite. The estimates are presented in Table 6, and the marginal correlation and cross-correlation fits for wind speed, temperature, and pressure are depicted in Figure 6 with dashed curves. The smoothness parameter is estimated as $\hat{\nu}_{22} = 3.30$ for temperature, which is much smoother than the estimate 1.32 from our flexible trivariate Matérn model, and the cross-correlation plot between wind and temperature in Figure 6 has obvious lack of fit. Table 7 compares the number of parameters in various models and the corresponding log-likelihood. Our flexible Matérn model achieves the highest log-likelihood with 21 parameters. Then more constraints lead to a smaller number of parameters in the model and a lower value of the log-likelihood.

Different predictive scores for the flexible and the parsimonious trivariate Matérn models are summarized in Table 8. The MSPE of our flexible trivariate Matérn model is 17.5, which is much smaller than the estimate 24.8, the MSPE of the parsimonious trivariate Matérn model. The other predictive scores, MAE, LogS, and CRPS, are all smaller for our model. Therefore, allowing for different scale parameters appears to be important for cokriging in this application. The prediction errors for each predicted location are shown in Figure 7.

Table 6. Maximum likelihood estimates of parameters for the parsimonious trivariate Matérn model applied to the wind speed, temperature, and pressure data. Units are miles per hour for wind speed, degree Celsius for temperature, Pascal for pressure, and kilometer for distances

$\hat{\nu}_{11}$	$\hat{\nu}_{22}$	$\hat{\nu}_{33}$	$\hat{\nu}_{12}$	$\hat{\nu}_{13}$	$\hat{\nu}_{23}$	$1/\hat{\alpha}$		
0.89	3.30	1.25	2.09	1.07	2.27	14.9		
$\hat{\sigma}_{11}$	$\hat{\sigma}_{22}$	$\hat{\sigma}_{33}$	$\hat{\rho}_{12}$	$\hat{\rho}_{13}$	$\hat{\rho}_{23}$	$\hat{\tau}_1$	$\hat{\tau}_2$	$\hat{\tau}_3$
11.72	1.23	0.15	0.43	−0.25	−0.29	0.25	1.04	0.20

6. DISCUSSION

We have introduced a valid parametric family of cross-covariance functions for multivariate spatial random fields where each component has a covariance function from a well-celebrated Matérn class. Specifically, we have presented conditions on the parameter space that result in valid models with varying degrees of complexity. Unlike the results of Gneiting et al. (2010) that were limited to the bivariate case, our model allows for various smoothnesses and rates of correlation decay for any number of components of the multivariate random field.

Table 7. Comparison of features of our flexible Matérn models with different parameter constraints and the parsimonious one for the wind speed, temperature, and pressure data

Model	Number of parameters	Loglik
Flexible Matérn	21	−34, 359.6
$\nu_{ij} = (\nu_{ii} + \nu_{jj})/2 + \Delta_A$ and $\alpha_{ij}^2 = (\alpha_{ii}^2 + \alpha_{jj}^2)/2 + \Delta_B$	17	−35, 125.6
$\nu_{ij} = (\nu_{ii} + \nu_{jj})/2 + \Delta_A$ and $\alpha_{ij}^2 = (\alpha_{ii}^2 + \alpha_{jj}^2)/2$	16	−35, 615.9
$\nu_{ij} = (\nu_{ii} + \nu_{jj})/2$ and $\alpha_{ij}^2 = (\alpha_{ii}^2 + \alpha_{jj}^2)/2 + \Delta_B$	16	−36, 193.3
Parsimonious Matérn	13	−36, 572.3

Table 8. Different predictive scores for wind speed with the flexible and parsimonious trivariate Matérn models

Model	MSPE	MAE	LogS	CRPS
Flexible	17.5	3.3	4.0	4.4
Parsimonious	24.8	3.8	4.4	5.3

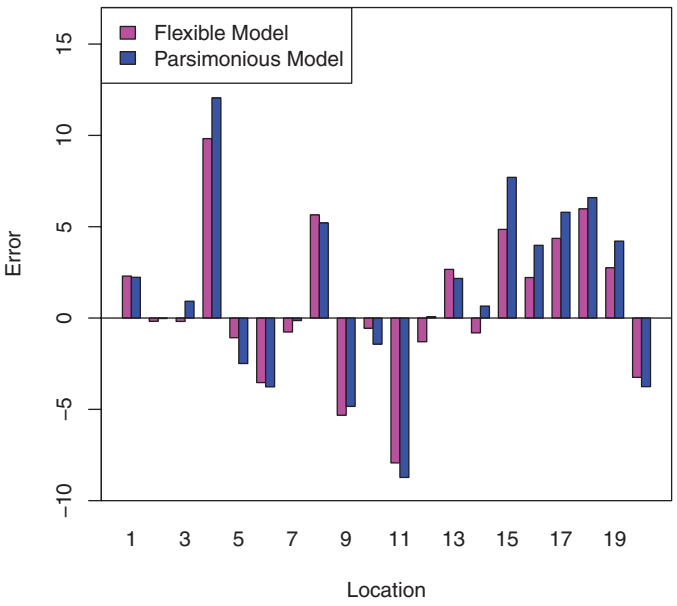

Figure 7. Prediction errors for the flexible (magenta) and parsimonious (blue) trivariate Matérn models for each of the 20 left-out locations. (The online version of this figure is in color.)

The formulation of Theorem 1 holds for any number of components, but imposes restrictions on the scale and smoothness parameters. Nevertheless, these conditions on our flexible multivariate Matérn model are far less restrictive than the ones proposed by Gneiting et al. (2010) for their parsimonious version. At the same time, the special case of our model for only two components compares well with the criterion proposed in Gneiting et al. (2010).

We have performed various Monte Carlo simulation experiments to demonstrate the performances of our approach. The application of the flexible multivariate Matérn model has been illustrated on a bivariate meteorological dataset of temperature/pressure over the Pacific Northwest that was previously analyzed by Gneiting et al. (2010). We have shown that our flexible bivariate Matérn model captured the physical features similarly to the optimal one from Gneiting et al. (2010). We also considered a trivariate dataset of wind/temperature/pressure in Oklahoma related to harvesting electricity from wind energy. We fitted our flexible Matérn model as well as the parsimonious model from Gneiting et al. (2010) to these data. We demonstrated that the prediction error associated with our flexible model was significantly smaller than that obtained with a parsimonious Matérn model. This suggests that the flexibility of multivariate Matérn models is important in modeling the cross-covariance functions of multivariate random fields with any number of components.

Apanasovich and Genton (2010) provided generic approaches to building asymmetric cross-covariance models. One can use such ideas for the proposed multivariate Matérn model, a topic for future research.

APPENDIX

Matrix Analysis Background

Let $\mathbf{A} = \{A_{ij}\}_{i,j=1}^{p}$ and $\mathbf{B} = \{B_{ij}\}_{i,j=1}^{p}$ be $p \times p$ matrices. The Hadamard product (or the entry-wise product) of $\mathbf{A}$ and $\mathbf{B}$ is the matrix $\mathbf{A} \circ \mathbf{B} = \{A_{ij} B_{ij}\}_{i,j=1}^{p}$. Issai Schur proved that if $\mathbf{A}$ and $\mathbf{B}$ are nonnegative definite, then so is $\mathbf{A} \circ \mathbf{B}$; see Horn and Johnson (1990). Suppose that $\mathbf{A}$ is nonnegative definite and that $A_{ij} \geq 0$ for all i and j. Then, $\mathbf{A}$ is *infinitely divisible* if the matrix $\mathbf{A}^{\circ r} = \{A_{ij}^{r}\}_{i,j=1}^{p}$ is nonnegative definite for every nonnegative r. It is easy to observe that every 2×2 nonnegative definite matrix with nonnegative entries is infinitely divisible. A Hermitian matrix $\mathbf{C}$ is *conditionally nonnegative definite* if $\mathbf{x}^{\mathrm{T}}\mathbf{C}\mathbf{x}^{*} \geq 0$, for all $\mathbf{x} \in \mathbb{C}^{p}$ such that $\sum_{i=1}^{p} x_i = 0$ and where $\mathbf{x}^{*}$ is the complex conjugate of $\mathbf{x}$. Charles Loewner showed that $\mathbf{C}$ is conditionally nonnegative definite and Hermitian if and only if the matrix $\{\exp(C_{ij})\}_{i,j=1}^{p}$ is infinitely divisible; see Guichardet (1972).

Lemmas

Lemma 1. Let $0 < b_i < \infty$, $i = 1, \ldots, m$, $\omega \geq 0$, and $\{B_{ij}\}_{i,j=1}^{p} > 0$ such that $\{\exp(-B_{ij})\}_{i,j=1}^{p}$ is infinitely divisible. Then, the $p \times p$ matrix $\mathbf{A}$ with entries

$$A_{ij} = \frac{\Gamma(b_i + b_j + \omega)}{B_{ij}^{b_i + b_j + \omega}}, \quad i, j = 1, \ldots, p,$$

is nonnegative definite.

Proof. It is clear that

$$\frac{\Gamma(b_i + b_j + \omega)}{B_{ij}^{b_i + b_j + \omega}} = \int_0^{\infty} \exp(-t B_{ij}) t^{(b_i + b_j + \omega) - 1} dt, \quad i, j = 1, \ldots, p.$$

For each $t \in (\epsilon, \infty)$, for some $\epsilon > 0$, the matrix with entries $\exp\{-t B_{ij}\} t^{(b_i + b_j + \omega) - 1}$ is nonnegative definite, then so is $\int_{\epsilon}^{\infty} \exp\{-t B_{ij}\} t^{(b_i + b_j + \omega) - 1} dt < \infty$. Next, take the limit for $\epsilon \to 0$. The claim follows. ■

Lemma 2. Let $\omega \geq 0$ and B_{ij}, $i, j = 1, \ldots, p$ be as in Lemma 1. Then, $\{(B_{ij} + \omega)^{-r}\}_{i,j=1}^{p}$ is nonnegative definite for any $0 < r < \infty$.

Proof. The claim follows from

$$\frac{1}{(B_{ij} + \omega)^{r}} = \int_0^{\infty} \exp(-t B_{ij}) \exp(-t\omega) t^{r-1} dt / \Gamma(r), \quad i, j = 1, \ldots, p.$$

■

Main Proofs

Proof of Theorem 1. First, assume that $\Delta_A > 0$. Following the statement of the theorem, let $A_{ij} = 1 - \{\nu_{ij} - (\nu_{ii} + \nu_{jj})/2\}/\Delta_A$, $V_{ij} = \sigma_{ij} \alpha_{ij}^{2\Delta_A + (\nu_{ii} + \nu_{jj})} \Gamma(\nu_{ij} + d/2) / [\Gamma\{(\nu_{ii} + \nu_{jj})/2 + d/2\} \Gamma(\nu_{ij})]$ and denote by $\imath = \sqrt{-1}$. It is known that $C_{ij}(\mathbf{h})$ is expressible in the form $C_{ij}(\mathbf{h}) = \int f_{ij}(\boldsymbol{\omega}) \exp(\imath \mathbf{h}^{\mathrm{T}} \boldsymbol{\omega}) d\boldsymbol{\omega}$, where $f_{ij}(\boldsymbol{\omega})$ are such that $\int |f_{ij}(\boldsymbol{\omega})| d\boldsymbol{\omega} < \infty$, $i, j = 1, \ldots, p$ (Stein 1999). Hence, by Cramér's theorem (Cramér 1945), we need to show that $\mathbf{F}(\boldsymbol{\omega}) = \{f_{ij}(\boldsymbol{\omega})\}_{i,j=1}^{p}$ is a nonnegative definite matrix for almost all $\boldsymbol{\omega} \in \mathbb{R}^{d}$. The spectral density of (3) is

$$\begin{aligned} f_{ij}(\boldsymbol{\omega}) &= \frac{1}{(2\pi)^{d}} \int \exp(-\imath \mathbf{h}^{\mathrm{T}} \boldsymbol{\omega}) C_{ij}(\mathbf{h}) d\mathbf{h} \\ &= \frac{\sigma_{ij} \alpha_{ij}^{2\nu_{ij}}}{(\alpha_{ij}^{2} + ||\boldsymbol{\omega}||^{2})^{\nu_{ij} + d/2}} \frac{\Gamma(\nu_{ij} + d/2)}{\Gamma(\nu_{ij})} \\ &= \frac{\Gamma\{(\nu_{ii} + \nu_{jj})/2 + d/2\}}{(\alpha_{ij}^{2} + ||\boldsymbol{\omega}||^{2})^{(\nu_{ii} + \nu_{jj})/2 + d/2}} \left(\frac{\alpha_{ij}^{2}}{\alpha_{ij}^{2} + ||\boldsymbol{\omega}||^{2}} \right)^{-\Delta_A A_{ij}} \\ &\quad \times \frac{1}{(\alpha_{ij}^{2} + ||\boldsymbol{\omega}||^{2})^{\Delta_A}} V_{ij}. \end{aligned}$$

We look at each term in the product. Matrices with elements $\Gamma\{(\nu_{ii} + \nu_{jj})/2 + d/2\}/(\alpha_{ij}^{2} + ||\boldsymbol{\omega}||^{2})^{(\nu_{ii} + \nu_{jj})/2 + d/2}$ and $(\alpha_{ij}^{2} + ||\boldsymbol{\omega}||^{2})^{-\Delta_A}$ are

nonnegative definite by Lemma 1 and Lemma 2, respectively. $\{V_{ij}\}_{i,j=1}^{p}$ is nonnegative by (iii) and the matrix with elements

$$\left(\frac{\alpha_{ij}^2}{\alpha_{ij}^2+||\boldsymbol{\omega}||^2}\right)^{-\Delta_A A_{ij}} = \exp\left[\Delta_A A_{ij}\left\{-\ln\left(1-\frac{||\boldsymbol{\omega}||^2}{\alpha_{ij}^2+||\boldsymbol{\omega}||^2}\right)\right\}\right], \quad i,j=1,\ldots,p,$$

is nonnegative definite for any $\boldsymbol{\omega}$ by Lemma 2 and repeatedly using the fact that the entry-wise product of nonnegative definite matrices is a nonnegative definite matrix. From Cramér's theorem, the claim follows. If $\Delta_A = 0$, then $\nu_{ij} = (\nu_{ii}+\nu_{jj})/2$ and

$$f_{ij}(\boldsymbol{\omega}) = \frac{\Gamma\{(\nu_{ii}+\nu_{jj})/2+d/2\}}{(\alpha_{ij}^2+||\boldsymbol{\omega}||^2)^{(\nu_{ii}+\nu_{jj})/2+d/2}} V_{ij}^*,$$

$$V_{ij}^* = \sigma_{ij}\alpha_{ij}^{(\nu_{ii}+\nu_{jj})}/\Gamma\{(\nu_{ii}+\nu_{jj})/2\}, \quad i,j=1,\ldots,p,$$

and applying Lemma 1, the claim follows. ■

Proof of Remark 1. We start with (a). Using the definition of a nonnegative definite matrix, take $\mathbf{x}\in\mathbb{C}^p$ such that $\sum_{i=1}^p x_i = 0$. Then,

$$\sum_{ij} x_i\alpha_{ij}^2 x_j^* = \sum_{ij} x_i\left\{\alpha_{ii}^2(\tau+1/2)+\alpha_{jj}^2(\tau+1/2)-2\tau\alpha_{ii}^2\alpha_{jj}^2\right\}x_j^*$$

$$= \left\{\left(\sum_i x_i\right)\left(\sum_j \alpha_{jj}^2 x_j^*\right)+\left(\sum_i x_i^*\right)\left(\sum_j \alpha_{jj}^2 x_j\right)\right\} \times(1/2+\tau)-2\tau\left(\sum_i x_i\alpha_{ii}^2\right)\left(\sum_i x_i^*\alpha_{ii}^2\right)\le 0.$$

Part (c) can be proved similarly. For (b), without loss of generality, assume that $0<\alpha_{11}^2<\alpha_{22}^2<\cdots<\alpha_{pp}^2$. Then,

$$\sum_{ij} x_i\max(\alpha_{ii}^2,\alpha_{jj}^2)x_j^* = \sum_i\sum_{j\ge i}(x_i x_j^*+x_i^* x_j)\alpha_{jj}^2 - \sum_i x_i x_i^*\alpha_{ii}^2 \le 0.$$

■

Proof of Remark 4. From the proof of Theorem 1, it follows that $\{\Gamma\{(\nu_{ii}+\nu_{jj})/2\}/\alpha_{ij}^{2(\nu_{ii}+\nu_{jj})/2}\}_{i,j=1}^p$ and $\{1/\alpha_{ij}^2\}_{i,j=1}^p$ form nonnegative definite matrices and $\mathcal{B}(\nu_{ij},d/2)\le\mathcal{B}\{(\nu_{ii}+\nu_{jj})/2,d/2\}$. Hence, $\tau_{ij}^{(k)}\le 1$, $k=1,2,3$. ■

Proof of Corollary 1. Part (a) follows directly from Theorem 1. Next, we sketch the proof of part (b). Let $V_{ij}=\sigma_{ij}\alpha_{ij}^{-d}/\Gamma(\nu_{ij})$. Then, the spectral density becomes

$$f_{ij}(\boldsymbol{\omega}) = \frac{V_{ij}\Gamma\{(\nu_i+\nu_j)/2+d/2\}}{\left(1+||\boldsymbol{\omega}||^2/\alpha_{ij}^2\right)^{(\nu_i+\nu_j)/2+d/2}}, \quad i,j=1,\ldots,p,$$

and from Lemma 1, the claim follows. ■

Proof of Corollary 2. The spectral density of (a) is

$$f_{ij}(\boldsymbol{\omega}) = \frac{\sigma_{ij}}{(\alpha^2+||\boldsymbol{\omega}||^2)^{d/2}}\exp[-\nu_{ij}\ln\{(\alpha^2+||\boldsymbol{\omega}||^2)/(\alpha^2)\}] \times\frac{\Gamma(\nu_{ij}+d/2)}{\Gamma(\nu_{ij})}, \quad i,j=1,\ldots,p,$$

which form a nonnegative definite matrix for any $\boldsymbol{\omega}$, and the claim follows. ■

Proof of Corollary 4. The claims follow directly from the previous results, noting that for the case of $p=2$, $\Delta_A=\nu_{12}-(\nu_{11}+\nu_{22})/2$. ■

[Received December 2010. Revised August 2011.]

REFERENCES

Apanasovich, T. V., and Genton, M. G. (2010), "Cross-Covariance Functions for Multivariate Random Fields Based on Latent Dimensions," *Biometrika*, 97, 15–30. [180,184,191]

Apanasovich, T. V., Ruppert, D., Lupton, J. R., Popovic, N., Turner, N. D., Chapkin, R. S., and Carroll, R. J. (2008), "Aberrant Crypt Foci and Semiparametric Modelling of Correlated Binary Data," *Biometrics*, 64, 490–500. [184]

Byrd, R. H., Lu, P., Nocedal, P., and Zhu, C. (1995), "A Limited Memory Algorithm for Bound Constrained Optimization," *SIAM Journal on Scientific Computing*, 16, 1190–1208. [185]

Cramér, H. (1945), *Mathematical Methods of Statistics*, Uppsala: Almqvist & Wiksells. [191]

Furrer, R., and Genton, M. G. (2011), "Aggregation-Cokriging for Highly-Multivariate Spatial Data," *Biometrika*, 98, 615–631. [185]

Gaspari, G., and Cohn, S. E. (1999), "Construction of Correlation Functions in Two and Three Dimensions," *Quarterly Journal of the Royal Meteorological Society*, 125, 723–757. [180]

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. [180]

Genton, M. G., and Hering, A. S. (2007), "Blowing in the Wind," *Significance*, 4, 11–14. [188]

Gneiting, T., Balabdaoui, F., and Raftery, A. E. (2007), "Probabilistic Forecasts, Calibration and Sharpness," *Journal of the Royal Statistical Society,* Series B, 69, 243–268. [185]

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), eds. B. Finkenstaedt, L. Held, and V. Isham, Boca Raton, FL: Chapman & Hall/CRC Press, pp. 151–175. [181]

Gneiting, T., Kleiber, W., and Schlather, M. (2010), "Matérn Cross-Covariance Functions for Multivariate Random Fields," *Journal of the American Statistical Association*, 105, 1167–1177. [181,182,183,185,186,187,190,191]

Gneiting, T., and Raftery, A. E. (2007), "Strictly Proper Scoring Rules, Prediction, and Estimation," *Journal of the American Statistical Association*, 102, 359–378. [185]

Goulard, M., and Voltz, M. (1992), "Linear Coregionalization Model: Tools for Estimation and Choice of Cross-Variogram Matrix," *Mathematical Geology*, 24, 269–282. [180]

Guichardet, A. (1972), *Symmetric Hilbert Spaces and Related Topics*, New York: Springer. [191]

Guttorp, P., and Gneiting, T. (2006), "Studies in the History of Probability and Statistics XLIX: On the Matérn Correlation Family," *Biometrika*, 93, 989–995. [181]

Hering, A. S., and Genton, M. G. (2010), "Powering Up With Space-Time Wind Forecasting," *Journal of the American Statistical Association*, 105, 92–104. [188]

Horn, R. A., and Johnson, C. R. (1990), *Matrix Analysis*, Cambridge: Cambridge University Press. [191]

Lindsay, B. G. (1988), "Composite Likelihood Methods," in *Statistical Inference From Stochastic Processes*, ed. N. U. Prabhu, Providence, RI: American Mathematical Society, pp. 221–239. [184]

Majumdar, A., and Gelfand, A. E. (2007), "Multivariate Spatial Modeling Using Convolved Covariance Functions," *Mathematical Geology*, 39, 225–245. [180]

Mardia, K. V., and Goodall, C. R. (1993), "Spatial-Temporal Analysis of Multivariate Environmental Monitoring Data," in *Multivariate Environmental Statistics* (North-Holland Series in Statistics and Probability No. 6), eds. G. P. Patil and C. R. Rao, Amsterdam: North-Holland, pp. 347–386. [180]

Matérn, B. (1960), *Spatial Variation*, New York: Springer. [181]

Mateu, J., Porcu, E., and Gregori, P. (2008), "Recent Advances to Model Anisotropic Space-Time Data," *Statistical Methods and Applications*, 17, 209–223. [180]

North, G. R., Wang, J., and Genton, M. G. (2011), "Correlation Models for Temperature Fields," *Journal of Climate*, 24, 5850–5862. [187]

Nott, D. J., and Dunsmuir, W. T. M. (2002), "Estimation of Nonstationary Spatial Covariance Structure," *Biometrika*, 89, 819–829. [180]

Porcu, E., Mateu, J., and Bevilacqua, M. (2007), "Covariance Functions That Are Stationary or Nonstationary in Space and Stationary in Time," *Statistica Neerlandica*, 61, 358–382. [180]

Schmidt, A. M., and Gelfand, A. E. (2003), "A Bayesian Coregionalization Approach for Multivariate Pollutant Data," *Journal of Geophysical Research*, 108, 1–9. [180]

Stein, M. L. (1999), *Interpolation of Spatial Data: Some Theory for Kriging*, Berlin: Springer. [181,191]

Ver Hoef, J. M., and Barry, R. P. (1998), "Constructing and Fitting Models for Cokriging and Multivariable Spatial Prediction," *Journal of Statistical Planning and Inference*, 69, 275–294. [180]

Wackernagel, H. (2003), *Multivariate Geostatistics: An Introduction With Applications* (3rd ed.), Berlin: Springer. [180]

Zhang, H. (2004), "Inconsistent Estimation and Asymptotically Equal Interpolations in Modelbased Geostatistics," *Journal of the American Statistical Association*, 99, 250–261. [182]

Zhang, H. (2007), "Maximum-Likelihood Estimation for Multivariate Spatial Linear Coregionalization Models," *Environmetrics*, 18, 125–139. [180]

Zhang, H., and Wang, Y. (2010), "Kriging and Cross-Validation for Massive Spatial Data," *Environmetrics*, 21, 290–304. [185]