Published in final edited form as:

Environmetrics. 2017 November ; 28(7): . doi:10.1002/env.2467.

Functional inverted Wishart for Bayesian multivariate spatial modeling with application to regional climatology model data

L. L. Duan¹, R. D. Szczesniak², and X. Wang³

¹Department of Statistical Science, Duke University, P.O. Box 90251, Durham, NC 27708, U.S.A.

²Division of Biostatistics and Epidemiology, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue (MLC 5041), Cincinnati, OH 45229, U.S.A.

³Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH 45221, U.S.A.

Abstract

Modern environmental and climatological studies produce multiple outcomes at high spatial resolutions. Multivariate spatial modeling is an established means to quantify cross-correlation among outcomes. However, existing models typically suffer from poor computational efficiency and lack the flexibility to simultaneously estimate auto- and cross-covariance structures. In this article, we undertake a novel construction of covariance by utilizing spectral convolution and by imposing an inverted Wishart prior on the cross-correlation structure. The cross-correlation structure with this functional inverted Wishart prior flexibly accommodates not only positive but also weak or negative associations among outcomes while preserving spatial resolution. Furthermore, the proposed model is computationally efficient and produces easily interpretable results, including the individual autocovariances and full cross-correlation matrices, as well as a partial cross-correlation matrix reflecting the outcome correlation after excluding the effects caused by spatial convolution. The model is examined using simulated data sets under different scenarios. It is also applied to the data from the North American Regional Climate Change Assessment Program, examining long-term associations between surface outcomes for air temperature, pressure, humidity, and radiation, on the land area of the North American West Coast. Results and predictive performance are compared with findings from approaches using convolution only or coregionalization.

Keywords

flexible covariance convolution; frequency-wise independence; NARCCAP; spectral latent variables

Correspondence. X. Wang, Department of Mathematical Sciences, University of Cincinnati, Cincinnati, OH 45221, U.S.A. xia.wang@uc.edu.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

1 | INTRODUCTION

In modern environmental and climatological studies, multivariate spatial data typically consist of observations on multiple outcomes (p) over a common set of locations (n). With advanced technology, it is not unusual to collect high-resolution spatial data across a large region. One of the crucial modeling tasks involves specifying a valid, flexible cross-covariance (Kleiber, 2017). Various solutions have been successfully developed, such as the linear model of coregionalization (Gelfand, Schmidt, Banerjee, & Sirmans, 2004; Schmidt & Gelfand, 2003), kernel convolution (Ver Hoef, Cressie, & Barry, 2004), covariance convolution (Majumdar & Gelfand, 2007), and the use of latent dimensions (Apanasovich & Genton, 2010). See Genton and Kleiber (2015) for a thorough review. On the other hand, the existing approaches are arguably very intensive in computation, due to the large size of cross-covariance; as both n and p increase, there is a clear need for scalable methods for estimation along with flexible modeling assumptions.

The motivating scientific inquiry arises from the research of a regional climatological system. It is of interest to statistically quantify the relationship among the long-term climate effects. In this study, we focus on four important surface outcomes (temperature, pressure, humidity, and radiation) that are produced by the regional climate models (RCMs) from the North American Regional Climate Change Assessment Program (NARCCAP; Mearns et al., 2011). This program, in the context of regional climate modeling and output, has been described in Neeley, Christensen, and Sain (2014). We restrict our focus to system outcomes for the land area of the West Coast in North America. Although the region is chosen to be small for approximate stationarity, due to the fine resolution ($50 \text{ km} \times 50 \text{ km}$), the location set remains sizable with $n = 2,870$. Even with only four outcomes, the covariance matrix already has a size of $11,480 \times 11,480$, which poses a heavy computational burden; meanwhile, $4 \cdot (4 - 1)/2 = 6$ long-term pairwise relationships need to be understood with a flexible model.

The popular standard solution for modeling such data is to use a Gaussian process with a parametric covariance matrix (Mardia & Goodall, 1993). For the multivariate data, the covariance matrix consists of autocovariances for each outcome on the diagonal blocks and the cross-covariance for each pair of outcomes on the off-diagonal blocks. To be a valid covariance, the matrix needs to be positive definite; therefore, it is important to constrain the cross-covariance. Among the early work, Apanasovich and Genton (2010) proposed augmenting the location space with an additional dimension and modeling the cross-correlation as a function decaying in the distance in this dimension; Majumdar and Gelfand (2007) used covariance convolution to combine two autocovariances to form a valid cross-covariance. These approaches meet the positive definiteness requirement, but are overly restrictive. The cross-covariance between any two outcomes can only be positive, yet in reality, the correlation can also be weak or negative. Kleiber (2017) illustrates the limitation from a spectral view in detail.

Another important challenge is the computational efficiency. The covariance matrix between np data points has $(np)^2$ elements, and the computational cost scales up quickly. A simple solution is using a Kronecker product of two matrices, one of $p \times p$ for interoutcome

correlation and the other of $n \times n$ for spatial correlation; however, this unrealistically assumes that the spatial correlation is the same for all outcomes. There are new covariances proposed for improving flexibility (Kleiber, 2017; Majumdar, Paul, & Bautista, 2010); to the best of our knowledge, they are limited for small data sets. As an alternative, coregionalization (Schmidt & Gelfand, 2003) is more widely used. The idea is to view p outcomes as a linear combination of r independent latent processes. This avoids the complication of $p(p-1)/2$ cross-covariances. However, due to the additional linear assumption, the drawback is that the modeling of each outcome is less flexible than a simple univariate marginal model. Furthermore, the autocovariance is a quadratic transform of r latent matrices, which is less straightforward than a simple parametric covariance.

These challenges motivate us to propose a flexible, efficient statistical approach for regional climate data. We combine the spectral convolution and a small $p \times p$ correlation matrix to induce a family of new flexible cross-covariances. The efficiency of spectral methods has been widely shown in univariate models (Duan, Wang, & Szczesniak, 2015; Fuentes, 2007), but underdeveloped for multivariate data. In the proposed method, each outcome has a closed-form autocovariance matrix formulated the same as in the univariate spatial model; jointly, the cross-covariance can accommodate different strengths of correlation (e.g., positive, weak, or negative). We further propose a latent variable approach for efficient posterior sampling, which completely avoids the direct computation of the $np \times np$ covariance. The high efficiency allows us to carry out full Bayesian estimation via Markov chain Monte Carlo (MCMC) and obtain posterior samples within a few hours. Conditionally on the latent variables, the $p \times p$ matrix follows a closed-form inverted Wishart distribution. We therefore name our method as “functional inverted Wishart” (FIW).

We organize this article as follows. In Section 2, we describe our model and its estimation procedure. We carry out a simulation study in Section 3 to show model performance in three different scenarios. In Section 4, we conduct analyses on the NARCCAP data with both exploratory analysis and the proposed model. In Section 5, we conclude with a discussion on result implications and potential extensions. The data set and the R codes are provided as Supporting Information.

2 | METHOD: FIW

2.1 | Notation and background

We use plain font for matrices and scalar quantities, and bold font for vectors. Let $S = \{s_i\}_{i=1, \dots, n} \in \mathbb{R}^d$ be a set of n -element locations in d -dimensional Euclidean space; at each location s_i , a p -element vector is observed $[Z_1(s_i), \dots, Z_p(s_i)]$, where $Z_k(s_i) \in \mathbb{R}$ for $k = 1, \dots, p$. Following common notation (Banerjee, Gelfand & Carlin, 2004), we use the functional form $Z_k(s_i) : \mathbb{R}^d \rightarrow \mathbb{R}$ because location s_i can be viewed as input and the observation as output. Denote the k th outcome over the whole space S by an n -dimensional column vector $\mathbf{Z}_k = [Z_k(s_1), Z_k(s_2), \dots, Z_k(s_n)]^T$ and all p outcomes over S by an np -dimensional column vector $\mathbf{Z} = [\mathbf{Z}_1^T, \dots, \mathbf{Z}_p^T]^T$. By a Gaussian assumption on $\mathbf{Z}$, the analysis of the multivariate outcomes $\mathbf{Z}$ focuses on the construction of the $np \times np$ covariance matrix $\mathbf{C}$. Let $C_{k,l}(s_i, s_j)$ represent the covariance between the k th outcome at location s_i and the l th

outcome at location s_j with $C_{k,\ell}(s_i, s_j) = \mathbb{E}[\{Z_k(s_i) - \mu_k(s_i)\}\{Z_\ell(s_j) - \mu_\ell(s_j)\}]$, where $\mu_k(s_i) = \mathbb{E}\{Z_k(s_i)\}$, $\mu_\ell(s_j) = \mathbb{E}\{Z_\ell(s_j)\}$, $k, \ell = 1, \dots, p$ and $s_i, s_j \in S$. This term is further defined as autocovariance of $Z_k(s_i)$ when $k = \ell$ and cross-covariance between $Z_k(s_i)$ and $Z_\ell(s_j)$ when $k \neq \ell$. For simplicity, we assume $\mu_k(s_i) = 0$ for $k = 1, \dots, p$ and all s_i as well as weak stationarity that $C_{k,\ell}(s_i, s_j) = C_{k,\ell}(s)$ with $s = s_i - s_j$. Although generalization to nonstationarity is possible, we focus on the stationary case for simplicity of illustration.

2.2 | FIW

We now define a new class of covariance functions by convolving two covariances and multiplying a scalar as follows:

$$C_{k,\ell}(\Delta s) = \rho_{k,\ell} \int_{h \in \mathbb{R}^d} C'_k(\Delta s) C'_\ell(\Delta s - h) dh + \delta_{k,\ell} \sigma_k^2, \quad (1)$$

where $\rho_{k,\ell}$ form a $p \times p$ positive definite matrix R with $\rho_{k,\ell}$ as its element on row k and column ℓ , $\rho_{k,\ell} = \rho_{\ell,k}$; $C'_k(\Delta s) = \int_{\omega \in \mathbb{R}^d} \exp\{i(\Delta s)^T \omega\} g_k^{1/2}(\omega) d\omega$, where the positive and finite function $g_k(\omega)$ is set as the spectral density of a known stationary covariance $C_k^*(\Delta s)$, $g_k(\omega) = \int_{\Delta s \in \mathbb{R}^d} \exp\{-i(\Delta s)^T \omega\} C_k^*(\Delta s) d\Delta s$. If $k = \ell$, $\delta_{k,\ell} = 1$, and if $k \neq \ell$, $\delta_{k,\ell} = 0$; $\sigma_k^2 \geq 0$ is the variance of noises, such as the measurement errors, for $k = 1, \dots, p$.

Using the convolution theorem, Equation 1 can be written as

$$C_{k,\ell}(\Delta s) = \rho_{k,\ell} \int_{\omega \in \mathbb{R}^d} \exp\{i(\Delta s)^T \omega\} g_k^{1/2}(\omega) g_\ell^{1/2}(\omega) d\omega + \delta_{k,\ell} \sigma_k^2. \quad (2)$$

One advantage of our parameterization in Equation 2 is that, despite the convolution, the autocovariance has a closed form $C_{k,k}(\Delta s) = \rho_{k,k} C_k^*(\Delta s) + \sigma_k^2$. This is different from previous convolution models, such as Majumdar and Gelfand (2007), which used $C'_k(\Delta s) = \int_{\omega \in \mathbb{R}^d} \exp\{i(\Delta s)^T \omega\} g_k(\omega) d\omega$ without halving the power of $g_k(\omega)$, leading to the autocovariance function $C_{k,k}(\Delta s) = \rho_{k,k} \int_{\omega \in \mathbb{R}^d} \exp\{i(\Delta s)^T \omega\} g_k(\omega) g_k(\omega) d\omega + \sigma_k^2$ that often lacks a closed form. This is an appealing property as it allows embedding a simple univariate spatial model marginally inside a multivariate one. This gives a simple interpretation for each marginal outcome under our model: the covariance parameters (such as the range and smoothness) in C_k^* have the same meanings as in the corresponding marginal spatial model.

Another advantage is that extra parameters $\rho_{k,\ell}$ with the associated R matrix significantly improves the flexibility of the cross-covariance structure. Consider the following cross-correlation $\rho(s)$ at distance s between two outcomes (excluding the variance of the independent noises σ^2):

$$\phi(\Delta s) = \frac{C_{k,\ell}(\Delta s) - \delta_{k,\ell} \sigma_k^2}{\sqrt{(C_{k,k}(\mathbf{0}) - \sigma_k^2)(C_{\ell,\ell}(\mathbf{0}) - \sigma_\ell^2)}} = \frac{\rho_{k,\ell}}{\sqrt{\rho_{k,k}\rho_{\ell,\ell}}} \cdot \frac{\int_{\mathbb{R}^d} C'_k(\Delta s) C'_\ell(\Delta s - \mathbf{h}) d\mathbf{h}}{\sqrt{\int_{\mathbb{R}^d} C'_k(\mathbf{0}) C'_k(\mathbf{h}) d\mathbf{h} \int_{\mathbb{R}^d} C'_\ell(\mathbf{0}) C'_\ell(\mathbf{h}) d\mathbf{h}}}. \quad (3)$$

The second term in Equation 3 is the effect caused by spatial convolution. We denote the first term in Equation 3 as

$$\tilde{\rho}_{k,\ell} = \frac{\rho_{k,\ell}}{\sqrt{\rho_{k,k}\rho_{\ell,\ell}}} \quad (4)$$

and interpret it as the “partial correlation” with the spatial convolution effects excluded. As discussed at the end of this section, the decomposition of the original cross-correlation $\phi(s)$ in Equation 3 reveals $\tilde{\rho}_{k,\ell}$ as a measure for the correlation between two outcomes in the frequency space.

It is clear in Equation 2 that the convolution combines the two stationary autocovariances to form a valid cross-covariance; however, using the covariance convolution alone would lead to a very restrictive model because it equates to implicitly assuming $\rho_{k,\ell} = 1$ among all outcomes. Figure 1 illustrates and compares the covariance structure with and without the free $\rho_{k,\ell}$ in the matrix R .

Without free $\rho_{k,\ell}$ in Equation 2,

$C_{k,\ell}(\Delta s) = \int_{\omega \in \mathbb{R}^d} \exp\{i(\Delta s)^T \omega\} g_k^{1/2}(\omega) g_\ell^{1/2}(\omega) d\omega + \delta_{k,\ell} \sigma_k^2$, which is a function of $g_k(\omega)$, $g_\ell(\omega)$, σ_k^2 and σ_ℓ^2 . As a result, once the spectral densities $g_k(\omega)$ and $g_\ell(\omega)$ are specified, that is, the autocovariances of outcomes k and ℓ are chosen, $\phi(s)$ would be completely deterministic at a given s . For example, with two autocovariances formed by the squared exponential structure with a unit scale parameter and the range parameters as 10 and 3, respectively, the cross-correlation $\phi(\mathbf{0})$ is fixed at a value of 0.741. This example is illustrated in the left panel of Figure 1 with the whole convolved cross-covariance function, whose values are fixed given the two autocovariances and are always positive. In comparison, with free $\rho_{k,\ell}$, the cross-covariance based on the same two autocovariances can vary from positive to negative values. This corresponds to the more realistic setting of the correlation between any two outcomes. The right panel of Figure 1 illustrates the added

flexibility, using $R = \begin{bmatrix} 1 & \rho_{1,2} \\ \rho_{1,2} & 1 \end{bmatrix}$ with various values for $\rho_{1,2}$.

From a theoretic standpoint, when all $C_k^*(\Delta s)$ and R are positive definite, the matrix formed by $C_{k,\ell}(s)$ in Equation 1 is positive definite. The proof is trivial by using the result from Majumdar et al. (2010). Further, Majumdar et al. (2010) suggested a spectrally varying $\rho_{k,\ell}(\omega)$ into the convolution integral, which is named as “coherence” by Kleiber (2017). We do

not adopt this extension as the majority of $g(\omega)$ is close to 0, for which the estimation for $\rho_{k,\ell}(\omega)$ is intractable, and $\rho_{k,\ell}(\omega)$ s lack a closed-form interpretation in the observation space.

In our model, assuming constant $\rho_{k,\ell}$ over ω not only allows borrowing information across frequencies but also makes it easy to interpret as a measure of the correlation between two outcomes in the frequency space. Using Figure 1 as an example, assume that the cross-correlation $\rho(0)$ in the observation space is found to be 0.73. As discussed earlier, the convolution alone would generate a correlation of 0.741 at $s = 0$, leading to a partial correlation of $\tilde{\rho}_{k,\ell} = 0.73/0.741 = 0.985$. This suggests that, although two outcomes may not seem heavily correlated in the observation space, they are in fact almost identical in the spectral space except differences in the spectral density. That is, in a spectral form

$Z_k(s_i) = \int_{\omega \in \mathbb{R}^d} \exp\{i(s_i)^T \omega\} g_k^{1/2}(\omega) U_k(\omega) d\omega + \varepsilon_k(s_i)$ with $\varepsilon_k \sim N(0, \sigma_k^2)$ for $k = 1, 2$, U_1 and U_2 are nearly identical even though Z_1 and Z_2 appear less similar due to discrepancy in their autocovariance parameters.

2.3 | Covariance approximation and posterior sampling

Because the covariance matrix is massive, it is almost computationally impossible to use it directly in the likelihood. In a univariate model, spectrally constructed covariances can be efficiently approximated by a discrete sum (Fuentes, 2007). Inspired by this idea, we first extend it to a multivariate covariance. Consider a discrete sum approximation to the integral in Equation 2 as follows:

$$\tilde{C}_{k,\ell}(\Delta s) = \rho_{k,\ell} \sum_{\omega_h \in \mathbf{W}} \exp\{i(\Delta s)^T \omega_h\} g_k^{1/2}(\omega_h) g_\ell^{1/2}(\omega_h) / m + \delta_{k,\ell} \sigma_k^2, \quad (5)$$

where ω_h are the members of the set

$$\mathbf{W} = \left\{ -\frac{m_1/2 - 1}{m_1/2} \pi, -\frac{m_1/2 - 2}{m_1/2} \pi, \dots, \frac{m_1/2 - 1}{m_1/2} \pi \right\} \times \dots \times \left\{ -\frac{m_d/2 - 1}{m_d/2} \pi, -\frac{m_d/2 - 2}{m_d/2} \pi, \dots, \frac{m_d/2 - 1}{m_d/2} \pi \right\}$$

, $m = m_1 m_2 \dots m_d$. Let $m_* := \min_{j=1, \dots, d} m_j$ if $\int_{\omega \notin (-\pi, \pi)^d} g_k^{1/2}(\omega) = O(1/m_*)$ for all k ; then, $\|C_{k,\ell}(\Delta s) - \tilde{C}_{k,\ell}(\Delta s)\| = O(1/m_*^2)$.

Let Q be the $n \times m$ factor matrix generated by $\{\exp(is_j^T \omega_h) / \sqrt{m}\}, j = 1, \dots, n$ and $h = 1, \dots, m$, Q^* be its conjugate transpose, G_k be the diagonal matrix with $g_k(\omega_h)$ as the h th diagonal element, and $R^{1/2}$ be the $p \times p$ matrix such that $(R^{1/2})^T R^{1/2} = R$. As $s = s_j - s_{j'}$, the covariance formed by Equation 5 can be written as a large matrix with $p \times p$ blocks, with the (k, ℓ) th block as an $n \times n$ matrix $Q G_k^{1/2} G_\ell^{1/2} Q^*$. The multivariate normal distribution specified for $\mathbf{Z}$ becomes

$$\begin{bmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \\ \vdots \\ \mathbf{Z}_p \end{bmatrix} \sim N \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \rho_{1,1} \mathbf{Q} \mathbf{G}_1^{1/2} \mathbf{G}_1^{1/2} \mathbf{Q}^* + I \sigma_1^2 & \rho_{1,2} \mathbf{Q} \mathbf{G}_1^{1/2} \mathbf{G}_2^{1/2} \mathbf{Q}^* & \cdots & \rho_{1,p} \mathbf{Q} \mathbf{G}_1^{1/2} \mathbf{G}_p^{1/2} \mathbf{Q}^* \\ \rho_{2,1} \mathbf{Q} \mathbf{G}_2^{1/2} \mathbf{G}_1^{1/2} \mathbf{Q}^* & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{p,1} \mathbf{Q} \mathbf{G}_p^{1/2} \mathbf{G}_1^{1/2} \mathbf{Q}^* & \cdots & \cdots & \rho_{p,p} \mathbf{Q} \mathbf{G}_p^{1/2} \mathbf{G}_p^{1/2} \mathbf{Q}^* + I \sigma_p^2 \end{bmatrix} \right). \quad (6)$$

In the univariate case, Whittle's likelihood can simplify the inverse and determinant calculation of the covariance (Whittle, 1953), leading to an efficient computing complexity of $\mathcal{O}(n \log_2 n)$. However, due to the block structure in distribution (6) the analytical inverse of the covariance in distribution (6) is hard to obtain, and the estimation of the spectral densities becomes very challenging. Direct adaptations of Whittle's approximation on multivariate data, such as Rosen and Stoffer (2007), no longer work here, as these adaptations rely on the Cholesky decomposition and are much more computationally intensive of $\mathcal{O}(np)^3$.

Instead of the previous approach, we consider a data augmentation approach that avoids the need to handle the covariance in its matrix form. Consider p many random vectors $\mathbf{Y}_k$, each with m elements, specified as $\mathbf{Y}_k(\boldsymbol{\omega}_h) = iY_{k,\text{Im}}(\boldsymbol{\omega}_h) + Y_{k,\text{Re}}(\boldsymbol{\omega}_h)$ with $h = 1, \dots, m$. For $\boldsymbol{\omega}_h \in [0, \pi)^d$, each latent variable $Y_k(\boldsymbol{\omega}_h)$ follows a complex normal distribution with

$Y_{k,\text{Im}} \stackrel{\text{iid}}{\sim} N(0, 1)$, $Y_{k,\text{Re}} \stackrel{\text{iid}}{\sim} N(0, 1)$; for $\boldsymbol{\omega}_h \notin [0, \pi)^d$, each $Y_k(\boldsymbol{\omega}_h)$ is determined by spectral symmetry $Y_k(\dots, \omega_a, \dots) = Y_k^*(\dots, -\omega_a, \dots)$. This symmetry ensures the cancelation of the imaginary i so that the covariance is real.

Let $\mathbf{Y} = [\mathbf{Y}_1 \| \mathbf{Y}_2 \| \dots \| \mathbf{Y}_p]$ be the juxtaposed $m \times p$ matrix. Multiplying $\mathbf{Y}$ by $R^{1/2}$ incorporates $\rho_{k,l}$ at each frequency and generates an $m \times p$ matrix $\mathbf{U} = \mathbf{Y} R^{1/2}$. Denote $\mathbf{U}_k$ as the k th column and $\mathbf{U}(\boldsymbol{\omega}_h)$ as the h th row of $\mathbf{U}$. With the $\mathbf{U}$ augmentation, the likelihood is significantly simplified because now, for $k = 1, \dots, p$, $h = 1, \dots, m$, we have

$\mathbf{Z}_k | \mathbf{U}_k \stackrel{\text{indep}}{\sim} N(\mathbf{Q} \mathbf{G}_k^{1/2} \mathbf{U}_k, \sigma_k^2 I)$, $\mathbf{U}_{\text{Re}}(\boldsymbol{\omega}_h) \stackrel{\text{iid}}{\sim} N(0, R)$, and $\mathbf{U}_{\text{Im}}(\boldsymbol{\omega}_h) \stackrel{\text{iid}}{\sim} N(0, R)$. These results are also shown in Figure 2.

The augmented $-2 \cdot \log$ -likelihood is then (omitting constants)

$$\begin{aligned} & \sum_{k=1}^p \left(\left\| \mathbf{Z}_k - \mathbf{Q} \mathbf{G}_k^{1/2} \mathbf{U}_k \right\|_c^2 / \sigma_k^2 + n \log \sigma_k^2 \right) \\ & + \sum_{\boldsymbol{\omega}_h \in [0, \pi)^d} \left\{ \mathbf{U}_{\text{Re}}^T(\boldsymbol{\omega}_h) R^{-1} \mathbf{U}_{\text{Im}}(\boldsymbol{\omega}_h) + \mathbf{U}_{\text{Im}}^T(\boldsymbol{\omega}_h) R^{-1} \mathbf{U}_{\text{Re}}(\boldsymbol{\omega}_h) + 2 \log |R| \right\}, \end{aligned}$$

where $\|\cdot\|_c$ is the complex modulus. When $\mathbf{U}$ is marginalized, this representation is equivalent to the one with the massive covariance in distribution (6). Because the distributions of $\mathbf{Z}_k$ s simply become independent normal given the $\mathbf{U}_k$ s, the computation is very efficient.

Using the data augmentation technique (Tanner & Wong, 1987), the posterior samples are updated via MCMC by alternating in the following three steps iteratively:

1. Given R and G_k s, update U from the complex normal distribution.
2. Given U , update the covariance parameters in G_k with the Metropolis–Hastings algorithm.
3. Given U , update R from the inverted Wishart distribution.

We provide the details of the posterior distributions in the Supporting Information. With lattice embedding (Duan et al., 2015), the overall complexity is $\mathcal{O}(pm \log_2 m)$, where m is the size of data lattice.

Since data augmentation is used, one concern might be how the rate of convergence is affected by the large number of latent variables. Although m is potentially large, due to the sparsity of the diagonal spectral density matrix G_k , the effective number of nonnegligible numbers in $G_k^{1/2} U_k$ is quite small (in our case, only about 3% of $G_k^{1/2} U_k$ has a modulus larger than 0.001, whereas the rest are very close to 0). As a result, it does not appear to heavily impact the geometric convergence of the covariance parameters. Theoretically speaking, the rate of geometric convergence depends on the Bayesian fraction of missing information (Liu, 1994), which is related to the maximal autocorrelation of the parameter in the Markov chain (Liu, 2001). Analytical solutions often do not exist, especially when the posterior does not have a closed form; however, one can approximate this rate by the elementwise autocorrelation (ACF) in the Markov chain (Liu, 1994). Our results show low autocorrelation for the covariance parameters within the first 100 lags, which indicates that the Bayesian fraction of missing information is relatively low for this advance model and the chain converges at a reasonable rate. Application details in our case study are provided at the end of Section 4.2.

2.4 | Predictive distribution

Let $V = \text{diag} \{ \sigma_k^2 \}_{k=1, \dots, p} \otimes I$ be the nugget diagonal $pn \times pn$ matrix. Denote the covariance matrix between another set of locations $\mathbf{S}'$ and the observed locations $\mathbf{S}$ by $K(\mathbf{S}', \mathbf{S})$ and the nugget at $\mathbf{S}'$ by V^* . The spatial predictive distribution for interpolation at $\mathbf{S}'$ is commonly defined as

$$\mathbf{Z}(\mathbf{S}') \sim N(K(\mathbf{S}', \mathbf{S})\{K(\mathbf{S}, \mathbf{S}) + V\}^{-1} \mathbf{Z}(\mathbf{S}), K(\mathbf{S}', \mathbf{S}) - K(\mathbf{S}', \mathbf{S})\{K(\mathbf{S}, \mathbf{S}) + V\}^{-1} K^T(\mathbf{S}', \mathbf{S}) + V^*).$$

The smoothing distribution is the same, except that $\mathbf{S}' = \mathbf{S}$ and V^* are equals zero. The direct estimation is computationally intensive. However, conditioning on the spectral latent variables U_k , the computation becomes very efficient with

$$\mathbf{Z}_k(\mathbf{S}') | U_k \sim N \left(Q_{\mathbf{S}'} G_k^{1/2} U_k, \sigma_k^2 I \right),$$

where the spatial autocorrelation and cross-correlation are captured in the convolution and the samples of U_k s.

Another important predictive distribution in the analysis of these data is related to multivariate spatial regression (Gelfand, Diggle, Guttorm, & Fuentes, 2010). Specifically, given observations on $p - 1$ outcomes and the cross-correlation matrix, the task is to predict the distribution of the remaining outcome (e.g., $\mathbf{Z}_k(\mathbf{S})$). Conditioning on the latent variables, the regression predictive distribution for $\mathbf{Z}_k(\mathbf{S})$ is

$$\mathbf{Z}_k(\mathbf{S})|\mathbf{U}_k^*, G_k, \sigma_k^2 \sim N\left(Q_S G_k^{1/2} \mathbf{U}_k^*, \sigma_k^2 I\right),$$

where each latent variable in $\mathbf{U}_k^*$ is specified as

$$U_k^*(\boldsymbol{\omega}_h) \sim \text{SCN}\left(R_{k,(-k)} R_{(-k),(-k)}^{-1} U_{(-k)}(\boldsymbol{\omega}_h), R_{k,k} - R_{k,(-k)} R_{(-k),(-k)}^{-1} R_{(-k),k}\right),$$

where the terms subscripted by $(-k)$ are those with the k th outcome removed. The symbol SCN denotes the symmetric complex normal distribution, where $x(\boldsymbol{\omega}_h) \sim \text{SCN}(a, b)$ means that, for $\boldsymbol{\omega}_h \in [0, \pi)^d$, $x_{\text{Im}}(\boldsymbol{\omega}_h) \stackrel{\text{iid}}{\sim} N(a, b)$, $x_{\text{Re}}(\boldsymbol{\omega}_h) \stackrel{\text{iid}}{\sim} N(a, b)$; for $\boldsymbol{\omega}_h \notin [0, \pi)^d$, $x(\dots, \omega_a, \dots) = x_k^*(\dots, -\omega_a, \dots)$.

3 | SIMULATION STUDY

To illustrate the advantages of the proposed approach, we carry out a simulation study. We generate 900 location points on the square space $(0, 30)^2$ and, on top of which, simulate a two-outcome spatial process via three different ways. For each setting, we leave out the bivariate outcome on 100 locations and use interpolation to predict them. We compute two metrics during validation: continuous ranked probability score (CRPS; Gneiting & Raftery, 2007) between the predictive distribution and the true value and the root mean square error (RMSE) between the predicted mean and the true value. For each model, we use a uniform prior $U(0, 10)$ for the range parameter τ and $U(0, 20)$ for all the scale parameters σ^2 , $R \sim IW(I, p + 1)$ ($p = 2$ here), and $\lambda_{jk} \stackrel{\text{iid}}{\sim} N(0, 1)$ for all j, k in the linear coregionalization model (see displayed equation). To exclude the effect of different initializations, we ran MCMC for 3,000 iterations and discarded the first 1,000 as burn-ins. We observed very similar numbers of effective sample size per 1,000 iterations (~ 200) in each setting, possibly due to the continuous Gaussian assumption.

We first generate data using linear combination of two independent Gaussian processes. This corresponds to the true model commonly known as coregionalization (Schmidt & Gelfand, 2003) as follows:

$$Z_k(\mathbf{s}_i) = \sum_{j=1}^J \lambda_{jk} Y_j(\mathbf{s}_i) + \varepsilon_k(\mathbf{s}_i),$$

for $k = 1, 2$, with $J = 2$, $\varepsilon_k(\mathbf{s}_i) \sim N(0, \sigma_k^2)$ for all $i = 1, \dots, 900$. We use $\lambda_{11} = 0.5$, $\lambda_{21} = -0.3$, $\lambda_{12} = 0.1$, $\lambda_{22} = 0.5$, and $\sigma_1^2 = \sigma_2^2 = 0.01$; $Y_1(\mathbf{s}_i)$ and $Y_2(\mathbf{s}_i)$ are Gaussian processes with mean

0 and the covariance formed by squared exponential function $\exp\{-\|s_i - s_{i'}\|^2/(2\tau_k^2)\}$, with $\tau_1 = 5$ for $Y_1(s_j)$ and $\tau_2 = 3$ for $Y_2(s_j)$. We run both FIW and coregionalization models on these data. Figures 3(a)–(c) plot the data and the estimated Gaussian process means. Column 1 of Table 1 shows the cross-validation results. We also compute the Frobenius norm of the difference between the true cross-covariance matrix and the posterior average of the estimated matrices. The coregionalization method produces a difference of 29.98, whereas the FIW shows a difference of 34.50. Although the proposed FIW is not the true model for these data, its performance is very close to that of the coregionalization model.

We next simulate data with the FIW as the true model. We choose the same squared exponential covariances in the last setting and convolve them to form the cross-covariance, using $\rho_{11} = \rho_{22} = 1$, $\rho_{12} = \rho_{21} = -0.6$, and $\sigma_1^2 = \sigma_2^2 = 0.01$. Figures 3(d)–(f) plot the data and the estimated Gaussian process means. Column 2 of Table 1 shows the cross-validation results. The FIW shows better performance in terms of interpolation. The coregionalization method shows a Frobenius difference of 47.20 from the true cross-covariance, whereas the FIW shows a difference of 24.49. This shows that the nonlinear cross-covariance formed by convolution could be challenging for the linear coregionalization to approximate.

Finally, we generate outcomes using trigonometric functions $Z_1(s_x, s_y) = \sin(s_x/5) + \cos(s_y/3) + e_1(s_x, s_y)$ and $Z_2(s_x, s_y) = \sin(s_x) + \cos(s_y) + e_2(s_x, s_y)$, with $e_1(s_x, s_y), e_2(s_x, s_y) \sim N(0, 0.3^2)$. The cross-covariance is not explicitly introduced as in either coregionalization or FIW. Figures 3(g)–(i) and Column 3 of Table 1 show that the two models are very similar in the mean estimation and interpolation.

To clarify, we do not claim that the proposed FIW always outperforms the coregionalization model. Instead, we show that coregionalization is a better fit when the cross-covariance is in fact linearly formed. At the same time, the FIW can accommodate the data better when the cross-covariance is nonlinearly formed. Therefore, these two models are two complementary ways to form cross-covariance. In terms of computing efficiency, because of the independent latent variables used in the FIW, its computation is substantially more efficient (Table 1). As one only needs to invert the small R matrix, the FIW is about 400 times faster than the coregionalization model, which needs to invert p number of $n \times n$ matrices. We expect that coregionalization can be accelerated similarly if our latent variable approach can be adopted.

4 | CASE STUDY

We now return to the study of the climatological data produced by RCMs from the NARCCAP program. The RCM is one of the dynamical downscaling methods that generates information at a high resolution. Simulations from RCMs enhance the climate impact studies on a local and regional scale at a spatial resolution of $50 \text{ km} \times 50 \text{ km}$. We use the WRFG + CGCM3 model, which is the Weather Research and Forecasting regional model driven under the boundary condition of the Third Generation Coupled Global Climate Model.

4.1 | Exploratory data analysis

We select four outcome variables that are approximately stationary in the West Coast area. That is, via a visual inspection of the observed data on the map, the range of correlation and

degree of smoothness of the observed outcomes appear homogeneous across the region. Section 5 discusses the possibility to extend the proposed model for nonstationary data. These four outcomes from the NARCCAP simulations are related to long-term effects, including temperature, pressure, humidity, and radiation. We take the seasonal average over July–September in the calendar year of 2000. Each variable contains entries at 2,870 geological locations. Table 2 shows the description, the acronyms, and the first two moment estimates for these variables.

We preprocessed the data by first removing the naïve trends using a linear model with intercept and polynomials (up to the third order) for the longitude and latitude coordinates of each outcome variable. We then standardized the residuals and used them as the inputs for all subsequent analyses. These procedures are performed to improve the numerical stability of the algorithm and simplify prior specification for the scale parameters. Figure 4 plots those four processed outcomes. It is immediately observed that temperature and pressure are much less smooth than humidity and radiation. Also, temperature and pressure seem to have a strong positive correlation.

The naïve Pearson correlation plot (Figure 5) confirms this observation. Pressure and temperature are indeed positively correlated (0.780). Furthermore, humidity and radiation levels appear negatively correlated (−0.798), whereas temperature and radiation levels show very weak correlation (0.001).

4.2 | FIW modeling

Now, we use the proposed model to analyze the data. Let $s = \{s_x, s_y\}$ and $w = \{w_1, w_2\}$. We choose the anisotropic Matérn function with $\nu = 3/2$ for each autocovariance function as follows:

$$C_{k,k}(s, s') = \rho_{k,k} \left(1 + \frac{|s_x - s'_x|}{\tau_{k1}} \right) \exp \left(-\frac{|s_x - s'_x|}{\tau_{k1}} \right) \cdot \left(1 + \frac{|s_y - s'_y|}{\tau_{k2}} \right) \exp \left(-\frac{|s_y - s'_y|}{\tau_{k2}} \right), \quad (7)$$

which has a spectral density $g_k(\omega_1, \omega_2) = 16\rho_{k,k}\tau_{k1}\tau_{k2} / \{(1 + \tau_{k1}^2\omega_1^2)^2 \cdot (1 + \tau_{k2}^2\omega_2^2)^2\}$, $k = 1, 2, 3, 4$. The choice of $\nu = 3/2$ is based on its balance between roughness and smoothness (differentiability): Stein (1999) found that the squared exponential (Matérn infinity) has too much differentiability, while $\nu = 1/2$ is not smooth enough. We chose $\nu = 3/2$ as a reasonable balancing choice. Our spectral approximation works well for any covariance that has a closed-form spectral density. This means that all covariances of Matérn family would work in our model.

With convolution, the cross-covariance between outcomes k and ℓ ($k, \ell = 1, \dots, 4$) is given by

$$C_{k,\ell}(s, s') = 16\rho_{k,\ell}(\tau_{k1}\tau_{k2}\tau_{\ell1}\tau_{\ell2})^{1/2} \cdot \int_{\omega \in \mathbb{R}^2} \frac{\exp\{i(s_x - s'_x)\omega_1 + i(s_y - s'_y)\omega_2\}}{(1 + \tau_{k1}^2\omega_1^2)(1 + \tau_{k2}^2\omega_2^2)(1 + \tau_{\ell1}^2\omega_1^2)(1 + \tau_{\ell2}^2\omega_2^2)} d\omega.$$

We use weakly informative priors for these parameters. They include the uniform $U(0, 20)$ for the parameters σ_k^2 , the normal $N(0, 50^2)$ truncated within the interval $(0, 30)$ for the range parameters $\tau_{\cdot}$, and the inverted Wishart $IW(I, p + 1)$ for R ($p = 4$ here). The upper truncation boundaries for the uniform are chosen on the basis of the effective maximum range of the observed data for $\tau_{\cdot}$ and a reasonable range of noises for σ_k^2 , given that the average magnitude is close to 1 with standardization. The posterior is also examined to make sure that they do not converge to values near the boundary.

We ran MCMC for 30,000 iterations, discarded the first 10,000 iterations, and thinned for every 10 steps. With the high computational efficiency, the procedures took 3.5 hr to complete with a single-threaded R program on a quad-core i7 laptop with 8 GB of memory. To empirically assess the convergence, we plot the autocorrelation (ACF) plots of the parameters. Figure 6 illustrates the convergence of the range parameters.

4.3 | Results

One useful property of the proposed model is that the autocovariance is in the same form as that in the univariate marginal model. Therefore, we first examine the range parameter estimates τ_{k1} and τ_{k2} for the four outcomes. Table 3 shows their posterior means and 95% credible intervals.

The range parameters determine the decay rate of autocorrelations in the Gaussian process. Figure 7 shows the posterior means for these fitted estimates. The estimates are very close to the observed data in Figure 4. This is a nice result from the marginal flexibility of the model.

It is useful to distinguish the actual spatial similarity/difference among the outcomes from the spurious correlation caused by the similarity in autocovariance parameters. In our model, the estimates of the cross-correlation (see Table 4) can be decomposed as a product of two parts as shown in Equation 3: one part of contribution from convolutions and the other from the partial correlation, respectively. Figure 8(a) plots the former at $s = \mathbf{0}$. Due to the similarity of range parameters (see Table 3), we find that temperature and pressure, as well as humidity and radiation, are almost perfectly correlated; at the same time, the other pairs, such as pressure and humidity, also seem highly correlated. As aforementioned, this suggests that the convolution should be only used as a mechanism to ensure positive definiteness. Figure 8(b) shows the estimates of the partial correlation defined in Equation 4. The partial correlation is quite useful, as it shows the remaining cross-correlation after the convolution effects are removed. For example, it shows that the temperature and pressure indeed have high similarities; however, although humidity and radiation are inversely correlated, the correlation is not as strong as the Pearson correlation indicated. This is consistent with the visual inspection of the data (Figure 7).

Finally, we assess the model performance using cross-validation. As a comparison, we consider two comparator models: (a) a convolution-only model without the R matrix; (b) a coregionalization model $Z_k(\mathbf{s}_i) = \sum_{j=1}^4 \lambda_{jk} Y_j(\mathbf{s}_i) + \varepsilon_k(\mathbf{s}_i)$ with $\varepsilon_k(\mathbf{s}_i) \sim N(0, \sigma_k^2)$ and $Y_j(\mathbf{s}_i)$ s as latent Gaussian processes. Their covariances are chosen to be the same form of anisotropic Matérn covariance as in Equation 7.

We first evaluate the interpolation performance. We delete the four-dimensional outcomes from a random set of 500 locations, run the model, and use the predictive distribution to predict the data. We compute the CRPS and RMSE between the predicted and observed data. Table 5 shows the results. Due to the limitation on the cross-covariance, the convolution-only model is a poor fit to the data and results in large prediction error, whereas the coregionalization model has better fit to the data. That said, both models have sizable interpolation errors. The FIW model outperforms each of the comparator models.

We next carry out cross-validation via regression, by removing one of the four outcomes completely and utilizing the cross-covariance for prediction. Table 6 shows the results. Due to the high correlation between temperature and pressure, the prediction is satisfactory for all three models, whereas for humidity and radiation, the convolution-only model suffers from large prediction error due to inflexible estimates on the cross-correlation. The coregionalization model shows improved performance, relative to the convolution-only model, whereas the FIW remains the best in prediction. In summary, among the three models, the proposed FIW provides the best fit to the surface outcome data.

5 | DISCUSSION

We have proposed an efficient Bayesian multivariate spatial model using the inverted Wishart distribution to improve the flexibility of the convolution-based cross-covariance function for modeling a local and stationary random field. We have demonstrated its advantages to convolution-only and coregionalization models in a case study of regional climatology data. This FIW model enjoys marginal interpretation with its marginal equivalence to the univariate model and high computational efficiency using the latent variable technique.

In the application to regional climatology data from the North American West Coast, we restrict our focus on the stationary field as a reasonable approximation. While this case study illustrates the benefits of the spectral approach and the need for flexibility in the cross-covariance, other large-field settings would require that the stationarity assumption be relaxed. In such cases, our model can be adopted in a locally stationary framework, such as the treed Gaussian process (Gramacy & Lee, 2008), for greater flexibility. Hence, it would be possible to allow R in the FIW model to vary from region to region under this framework, similar to Gelfand et al. (2004), thereby achieving nonstationarity. Furthermore, some portions of the model motif, such as the R matrix, can include a dynamic structure to accommodate the temporal evolution in the data (West, 2014).

The latent variable approach that we proposed can be extended to estimate a spectrally varying coherence (Kleiber, 2017). Even though this would remove the ability to interpret

the R matrix in the observed space, it could have utility for those researchers interested in the spectral properties of the multivariate random field.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Apanasovich TV, Genton MG. Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika*. 2010; 97(1):15–30.
- Banerjee, S., Gelfand, AE., Carlin, BP. Hierarchical modeling and analysis for spatial data. Boca Raton, FL: Chapman Hall/CRC; 2004.
- Duan LL, Wang X, Szczesniak RD. Functional Gaussian process model for Bayesian nonparametric analysis. 2015 Retrieved from arXiv:1502.03042.
- Fuentes M. Approximate likelihood for large irregularly spaced spatial data. *Journal of the American Statistical Association*. 2007; 102(477):321–331. [PubMed: 19079638]
- Gelfand, AE., Diggle, P., Guttorp, P., Fuentes, M. Handbook of spatial statistics. Boca Raton, FL: Chapman Hall/CRC; 2010.
- Gelfand AE, Schmidt AM, Banerjee S, Sirmans C. Nonstationary multivariate process modeling through spatially varying coregionalization. *Test*. 2004; 13(2):263–312.
- Genton MG, Kleiber W. Cross-covariance functions for multivariate geostatistics. *Statistical Science*. 2015; 30(2):147–163.
- Gneiting T, Raftery AE. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*. 2007; 102(477):359–378.
- Gramacy RB, Lee HK. Bayesian treed Gaussian process models with an application to computer modeling. *Journal of the American Statistical Association*. 2008; 103(483):1119–1130.
- Kleiber W. Coherence for multivariate random fields. *Statistica Sinica*. 2017 In press.
- Liu, JS. The fraction of missing information and convergence rate for data augmentation. *Computing Science and Statistics*. In: Sall, J., Lehman, A., editors. *Computing Science and Statistics: Proceedings of the 26th symposium on the Interface*. Fairfax Station, VA: Interface Foundation of North America; 1994. p. 490–496.
- Liu, JS. Monte Carlo strategies in scientific computing. New York: Springer; 2001.
- Majumdar A, Gelfand AE. Multivariate spatial modeling for geostatistical data using convolved covariance functions. *Mathematical Geology*. 2007; 39(2):225–245.
- Majumdar A, Paul D, Bautista D. A generalized convolution model for multivariate nonstationary spatial processes. *Statistica Sinica*. 2010; 20(2):675–695.
- Mardia KV, Goodall CR. Spatial-temporal analysis of multivariate environmental monitoring data. *Multivariate Environmental Statistics*. 1993; 6(76):347–385.
- Mearns, L., Gutowski, W., Jones, R., Leung, L., McGinnis, S., Nunes, A., Qian, Y. The North American Regional Climate Change Assessment Program dataset National Center for atmospheric research earth system grid data portal. Boulder, CO: 2011. Data downloaded: 01-03-2016
- Neeley E, Christensen W, Sain S. A Bayesian spatial factor analysis approach for combining climate model ensembles. *Environmetrics*. 2014; 25(7):483–497.
- Rosen O, Stoffer DS. Automatic estimation of multivariate spectra via smoothing splines. *Biometrika*. 2007; 94(2):335–345.
- Schmidt AM, Gelfand AE. A Bayesian coregionalization approach for multivariate pollutant data. *Journal of Geophysical Research: Atmospheres*. 2003; 108(D24):8783–8791.
- Stein, ML. Interpolation of spatial data: Some theory for kriging. New York: Springer; 1999.
- Tanner MA, Wong WH. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*. 1987; 82(398):528–540.

- Ver Hoef JM, Cressie N, Barry RP. Flexible spatial models for kriging and cokriging using moving averages and the fast Fourier transform (FFT). *Journal of Computational and Graphical Statistics*. 2004; 13(2):265–282.
- West, M. Bayesian forecasting. New York: Wiley; 2014. StatsRef: Statistics Reference Online
- Whittle P. The analysis of multiple stationary time series. *Journal of the Royal Statistical Society. Series B (Methodological)*. 1953; 15(1):125–139.

FIGURE 1.

Two squared exponential autocovariances (Auto-Cov 1 with range 10 and Auto-Cov 2 with range 3) are convolved to form a cross-covariance. Standard convolution without using the R matrix limits the flexibility of cross-correlation (left panel). For example, at distance 0, the cross-correlation is always fixed to 0.741; the FIW solves this problem and accommodates weak or negative cross-correlation (right panel)

FIGURE 2.

The latent variable view for the conditional mean of the data. Independent complex normal vectors Y_k are first multiplied by $R^{1/2}$ to form correlated spectral vectors U_k . Then, each U_k is multiplied by the half spectral density $G_k^{1/2}$ and Fourier transform matrix Q to form the conditional mean of each observation vector Z_k

(a) Simulated data from co-regionalization model. (b) Estimated mean using co-regionalization. (c) Estimated mean using FIW.

(d) Simulated data from FIW model. (e) Estimated mean using co-regionalization. (f) Estimated mean using FIW.

(g) Simulated data from trigonometric functions. (h) Estimated mean using co-regionalization. (i) Estimated mean using FIW.

FIGURE 3.

Data simulated from three models and the estimated Gaussian process means via coregionalization and functional inverted Wishart (FIW). The FIW is very similar to the coregionalization in terms of smoothing

FIGURE 4.

Four long-term effect outcomes: surface air temperature (tas), surface pressure (ps), surface specific humidity (huss), and surface downwelling shortwave radiation (rsds). The first two show strong similarity, and the last two appear inversely correlated

FIGURE 5.

Pearson correlation of the four outcomes (detrended and standardized), with the correlation values ranging from -1 (blue) to 1 (red). The data have different levels of correlation that need to be accommodated by a flexible cross-covariance function. tas = surface air temperature; ps = surface pressure; huss = surface specific humidity; rsds = surface downwelling shortwave radiation

FIGURE 6.

Autocorrelation plots show that the covariance parameters τ in the functional inverted Wishart rapidly converge within the first 100 lags

FIGURE 7.

Estimated Gaussian process means show that fine details are captured by the functional inverted Wishart. tas = surface air temperature; ps = surface pressure; huss = surface specific humidity; rsds = surface downwelling shortwave radiation

FIGURE 8.

Panel (a) shows the convolution resulting in strong and positive correlation for all pairs of outcomes; panel (b) shows the R matrix introduced in the functional inverted Wishart accommodating positive, weak, or negative correlations. The former ensures positive definiteness of the overall covariance matrix, and the latter allows a more realistic modeling of the data. The partial correlation matrix shown in (b) alone can be interpreted as the R matrix excluding the convolution effects. tas = surface air temperature; ps = surface pressure; huss = surface specific humidity; rsds = surface downwelling shortwave radiation.

TABLE 1

Cross-validation using interpolation in simulations

	Sim 1		Sim 2		Sim 3		Average seconds per 1,000 samplings $\pm$ <i>SD</i>
	CRPS	RMSE	CRPS	RMSE	CRPS	RMSE	
Coregionalization	0.063	0.117	0.093	0.150	0.231	0.427	36, 538 $\pm$ 1, 700
FIW	0.066	0.130	0.064	0.106	0.227	0.407	85 $\pm$ 10

Note. CRPS = continuous ranked probability score; RMSE = root mean square error; FIW = functional inverted Wishart.

Lower values of CRPS, RMSE, and computational time are bold to indicate the better model.

TABLE 2

Description and summary statistics for the four outcomes in the WRFG+CGCM3

Acronym	Description	Units	<i>M</i>	<i>SD</i>
tas	Surface air temperature	K	286.55	7.12
ps	Surface pressure	Pa	94133.09	6406.56
huss	Surface specific humidity	$10^{-3}\text{kg}\cdot\text{kg}^{-1}$	6.41	3.10
rsds	Surface downwelling shortwave radiation	$\text{W}\cdot\text{m}^{-2}$	230.96	58.71

TABLE 3

The range estimates with 95 % credible intervals for the covariance parameter

Outcome	τ_{k1}	τ_{k2}
tas	1.69 (1.62, 1.77)	2.39 (2.28, 2.49)
ps	1.85 (2.28, 2.49)	2.47 (2.33, 2.58)
huss	3.89 (3.60, 4.22)	5.39 (4.90, 5.89)
rsds	3.85 (3.42, 4.28)	4.57 (4.21, 5.02)

Note. The estimates can be interpreted marginally in the same way as in univariate marginal spatial models. The numbers are shown in 50 km/unit. tas = surface air temperature; ps = surface pressure; huss = surface specific humidity; rsds = surface downwelling shortwave radiation.

TABLE 4

The covariance estimate for the R matrix with 95% credible intervals

Outcome	tas	ps	huss	rsds
tas	0.287 (0.280, 0.301)	–	–	–
ps	0.304 (0.284, 0.315)	0.410 (0.371, 0.430)	–	–
huss	0.099 (0.092, 0.105)	0.198 (0.162, 0.225)	0.385 (0.327, 0.467)	–
rsds	0.025 (0.023, 0.026)	–0.022 (–0.049, 0.010)	–0.122 (–0.147, –0.096)	0.326 (0.289, 0.367)

Note. tas = surface air temperature; ps = surface pressure; huss = surface specific humidity; rsds = surface downwelling shortwave radiation.

TABLE 5

Cross-validation using interpolation

	tas		ps		huss		rsds	
	CRPS	RMSE	CRPS	RMSE	CRPS	RMSE	CRPS	RMSE
Convolution only	0.35	0.64	0.31	0.47	0.32	0.77	0.40	0.53
Coregionalization	0.33	0.59	0.29	0.41	0.17	0.36	0.25	0.41
FIW	0.18	0.25	0.14	0.26	0.08	0.14	0.09	0.13

Note. Compared with the convolution-only model without the R matrix and the coregionalization model, the functional inverted Wishart (FIW) has better fit to the data and lower prediction error measured by continued rank probability score (CRPS) and root mean square error (RMSE). tas = surface air temperature; ps = surface pressure; huss = surface specific humidity; rsds = surface downwelling shortwave radiation.

TABLE 6

Cross-validation using regression

	tas		ps		huss		rsds	
	CRPS	RMSE	CRPS	RMSE	CRPS	RMSE	CRPS	RMSE
Convolution only	0.13	0.25	0.20	0.34	0.78	1.08	0.59	0.93
Coregionalization	0.13	0.24	0.18	0.30	0.32	0.40	0.28	0.42
FIW	0.12	0.23	0.19	0.30	0.28	0.35	0.20	0.34

Note. Convolution-only model suffers from poor performance in regression due to the inflexibility in the cross-covariance, whereas coregionalization and functional inverted Wishart (FIW) models have substantially lower error measured by continued rank probability score (CRPS) and root mean square error (RMSE). tas = surface air temperature; ps = surface pressure; huss = surface specific humidity; rsds = surface downwelling shortwave radiation.