COHERENCE FOR MULTIVARIATE RANDOM FIELDS

William Kleiber

University of Colorado

Abstract: Multivariate spatial field data are increasingly common and their modeling typically relies on building cross-covariance functions to describe cross-process relationships. An alternative viewpoint is to model the matrix of spectral measures. We develop the notions of coherence, phase, and gain for multidimensional stationary processes. Coherence, as a function of frequency, is a measure of linear relationship between two spatial processes at that frequency. We use the coherence function to illustrate fundamental limitations on a number of previously proposed constructions for multivariate processes, suggesting these options are not viable for data modeling. We also give natural interpretations to cross-covariance parameters of the Matérn class, where the cross-smoothness controls the decay of coherence at infinitely high frequencies, and the cross-range parameter controls the frequency of greatest coherence. These interpretations provide warnings for particular parameter combinations that imply potentially non-physical relationships between variables. Estimation follows from smoothed multivariate periodogram matrices. We illustrate the estimation and interpretation of these functions on two datasets, forecast and reanalysis sea level pressure and geopotential heights over the equatorial region.

Key words and phrases: Coherency, gain, periodogram, phase, reanalysis, squared coherence, spectral density.

1. Introduction

The theory of univariate continuous stochastic processes has become well developed over nearly a century of research. The past quarter century or so has seen an increasing interest and development of models for multivariate spatial processes. The recent review by Genton and Kleiber (2015) gives a comprehensive treatment of the basic approaches that have been explored to build stochastic spatial models. In the discussion, Bevilacqua, Hering and Porcu (2015) posed the question of, given the recent deluge of multivariate constructions, "which parametric model is more flexible?" Indeed, the relative strengths and weaknesses of multivariate models have been explored only by testing a battery of different models on particular datasets, and comparing performance either by likelihood values or by predictive cross-validation (in the multivariate context, spatial prediction is known as co-kriging). Thus, a fundamental open question is (1), to

what extent can the flexibilities of model constructions be compared theoretically? Additionally, most models are motivated in the covariance domain, and the natural follow-up question is (2), are there other approaches than covariance to measure and quantify spatial dependence?

Spectral coherence has traditionally appeared in the time series literature as a quantitative method to describe the strength of relationship between components of a multivariate time series (Koopmans (1964); Brockwell and Davis (2009); Shumway and Stoffer (2011)). In the spatial context, coherence has been successfully applied in other sciences such as optics (Mandel and Wolf (1976); Carter and Wolf (1977)), while in the statistical literature there appears to be a less thorough history. For instance, Fuentes (2006) used coherence to test for time-separability of space-time fields, while Guinness et al. (2014) suggested its use for modeling conditional dependence between processes. We introduce and discuss the notions of spectral coherence, phase, and gain for multidimensional and multivariate spatial random fields. We propose that these functions allow for natural partial answers to questions (1) and (2).

From coherence, we illustrate fundamental limitations on a number of multivariate models in that they imply constant coherence across all frequencies. We also suggest insights into such well-established models as the multivariate Matérn, where parameters such as the cross-covariance smoothness and range have had elusive direct interpretations that relate to process behavior. Moreover, certain parameter settings imply possibly non-physical behavior, for instance implying coherence that does not decay to zero at arbitrarily high frequencies. For instance, self-similar processes such as multivariate fractional Brownian motion have constant coherence (Amblard et al. (2012)), whereas a number of empirical studies in the various sciences suggest coherence should decay at high frequencies (Fante (1974); Katz and Briscoe (1979); Kneer et al. (1980); Schreiner and Dorman (1990)), although the outcomes of some are less clear (Mack and Flinn (1971)).

We consider two datasets from the atmospheric sciences. Both are reforecast and reanalysis data products over the equatorial region based on a well established numerical weather prediction (NWP) model. Reanalysis forecasts are from a fixed version of a NWP model that are run retrospectively to generate a large database of model forecasts and analyses (in this context, an analysis can be considered a best estimate of the current state of the atmosphere). First we look at forecasted surfaces of sea level pressure at daily forecast horizons between 24 and 192 hours. We show that coherence can be used as a diagnostic to assess forecast

quality, and additionally illustrate frequency bands at which forecasts improve over time. The second dataset involves geopotential heights at differing pressure levels. We show that coherence and phase extract, and highlight qualities of the spatial relationship between different pressure levels that are difficult to model using extant multivariate covariance constructions, and indeed illustrate major limitations of existing popular constructions.

2. Spectra for Multivariate Random Fields

Suppose $\mathbf{Z}(\mathbf{s}) = (Z_1(\mathbf{s}), \ldots, Z_p(\mathbf{s}))^{\mathrm{T}} \in \mathbb{C}^p$ is a p-variate weakly stationary random field on $\mathbf{s} \in \mathbb{R}^d$ admitting a matrix-valued covariance function $\mathbf{C}(\mathbf{h}) = (C_{ij}(\mathbf{h}))_{i,j=1}^p$ where $C_{ij}(\mathbf{h}) = \mathrm{Cov}(Z_i(\mathbf{s} + \mathbf{h}), Z_j(\mathbf{s}))$. For simplicity of exposition we suppose $\mathbf{Z}(\mathbf{s})$ is a mean zero process. For complex-valued stationary processes, $\mathrm{Cov}(Z_i(\mathbf{s}_1), Z_j(\mathbf{s}_2)) = \mathbb{E}(Z_i(\mathbf{s}_1)\overline{Z_j(\mathbf{s}_2)})$, so that $C_{ij}(\mathbf{h}) = \overline{C_{ji}(\mathbf{h})}$. The main obstacle to multivariate process modeling is developing flexible classes of matrix-valued covariance functions $\mathbf{C}$ that are nonnegative definite.

For univariate processes, Bochner's Theorem states that $C_{ii}(\mathbf{h})$ is a valid (i.e., nonnegative definite) function if and only if it can be written

$$C(\mathbf{h}) = \int_{\mathbb{R}^d} \exp(i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{h})\mathrm{d}F(\boldsymbol{\omega}),$$

where F is a positive finite measure (Stein (1999)). If F admits a density f with respect to the Lebesgue measure on $\mathbb{R}^d$, we call it the spectral density for C. The multivariate extension of Bochner's fundamental result is given by Cramér (1940), here stated for covariances admitting spectral densities.

Theorem 1 (Cramér (1940)). *A matrix-valued function* $\mathbf{C} : \mathbb{R}^d \to \mathbb{C}^{p \times p}, \mathbf{C} = (C_{ij})_{i,j=1}^p$ *is nonnegative definite if and only if*

$$C_{ij}(\mathbf{h}) = \int_{\mathbb{R}^d} \exp(i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{h})f_{ij}(\boldsymbol{\omega})\mathrm{d}\boldsymbol{\omega}$$

for $i, j = 1, \ldots, p$ *such that the matrix* $\mathbf{f}(\boldsymbol{\omega}) = (f_{ij}(\boldsymbol{\omega}))_{i,j=1}^p$ *is nonnegative definite for all* $\boldsymbol{\omega} \in \mathbb{R}^d$.

The functions $f_{ij}(\boldsymbol{\omega})$ are the spectral and cross-spectral densities for the marginal and cross-covariance functions $C_{ij}(\mathbf{h})$, and $f_{ij}(\boldsymbol{\omega}) = \overline{f_{ji}(\boldsymbol{\omega})}$. When the spectral density exists, it can be solved for as the Fourier transform of the covariance function,

$$f_{ij}(\boldsymbol{\omega}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \exp(-i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{h})C_{ij}(\mathbf{h})\mathrm{d}\mathbf{h}.$$

Theorem 1 has primarily been used in practice to build multivariate covariance

models, by specifying matrices of spectral densities that are nonnegative definite for all frequencies.

2.1. Coherence

In time series, the notion of frequency coherence is well developed, and can be used, for instance, to assess whether one time series is related to another by a time invariant linear filter. These notions carry over to the spatial case, and form the point of entry for our analyses.

If $C_{ij}(\mathbf{h}), i, j = 1, 2$, form a matrix-valued covariance function with associated spectral densities $f_{ij}(\boldsymbol{\omega})$, then define the coherence function (or coherency function)

$$\gamma(\boldsymbol{\omega}) = \frac{f_{12}(\boldsymbol{\omega})}{\sqrt{f_{11}(\boldsymbol{\omega})f_{22}(\boldsymbol{\omega})}}.$$

We might assume $f_{ii}(\boldsymbol{\omega}) > 0$ for all $\boldsymbol{\omega} \in \mathbb{R}^d$ for $i = 1, 2$, but can define $\gamma(\boldsymbol{\omega}) = 0$ if $f_{ii}(\boldsymbol{\omega}) = 0$. The coherence function can be complex-valued, so in practice we examine the absolute coherence function, $|\gamma(\boldsymbol{\omega})|$. The real-valued function $|\gamma(\boldsymbol{\omega})|^2$ is the squared coherence function and, by Theorem 1, $0 \le |\gamma(\boldsymbol{\omega})|^2 \le 1$ for all $\boldsymbol{\omega}$. Values of $|\gamma(\boldsymbol{\omega})|$ near unity indicate a linear relationship between $Z_1(\mathbf{s})$ and $Z_2(\mathbf{s})$ at particular frequency bands.

The predictive estimator in the next result is based on a kernel-smoothed process, a natural predictor given the interpretation of the univariate kriging weights as a kernel function (Kleiber and Nychka (2015)).

Theorem 2. *Suppose* $(Z_1(\mathbf{s}), Z_2(\mathbf{s}))^{\mathrm{T}}$ *is a complex-valued mean zero weakly stationary bivariate field with matrix-valued covariance* $\mathbf{C}(\mathbf{h})$ *admitting a spectral density matrix* $\mathbf{f}(\boldsymbol{\omega}) = (f_{ij}(\boldsymbol{\omega}))_{i,j=1}^2$ *that is everywhere nonzero. Then the continuous square integrable function* $K(\mathbf{u}) : \mathbb{R}^d \to \mathbb{C}$ *that minimizes* $\mathbb{E}|Z_1(\mathbf{s}_0) - \int_{\mathbb{R}^d} K(\mathbf{u} - \mathbf{s}_0)Z_2(\mathbf{u})\mathrm{d}\mathbf{u}|^2$ *is*

$$K(\mathbf{u}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \exp(-i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{u}) \frac{f_{12}(\boldsymbol{\omega})}{f_{22}(\boldsymbol{\omega})} \mathrm{d}\boldsymbol{\omega}$$

$$= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \exp(-i\boldsymbol{\omega}^{\mathrm{T}}\mathbf{u}) \sqrt{\frac{f_{11}(\boldsymbol{\omega})}{f_{22}(\boldsymbol{\omega})}} \gamma(\boldsymbol{\omega}) \mathrm{d}\boldsymbol{\omega}. \tag{2.1}$$

The spectral density of the predictor $\hat{Z}_1(\mathbf{s}_0) = \int_{\mathbb{R}^d} K(\mathbf{u} - \mathbf{s}_0)Z_2(\mathbf{u})\mathrm{d}\mathbf{u}$ *is, for all* $\boldsymbol{\omega} \in \mathbb{R}^d$,

$$f_{1|2}(\boldsymbol{\omega}) = f_{11}(\boldsymbol{\omega})|\gamma(\boldsymbol{\omega})|^2. \tag{2.2}$$

The relationship (2.1) implies that the optimal weighting function is modu-

lated by the coherence between the two processes, and indeed has greater spectral weight on frequencies with high coherence. This can be contrasted with the analogous result for time series, see, e.g., Theorem 8.3.1 of Brillinger (2001).

An immediate corollary to Theorem 2 is that

$$|\gamma(\boldsymbol{\omega})|^2 = \frac{f_{1|2}(\boldsymbol{\omega})}{f_{11}(\boldsymbol{\omega})}.$$

Thus, the coherence has an attractive interpretation as the amount of variability that can be attributed to a linear relationship between two processes at a particular frequency. Reich, Chang and Foley (2014) uses a similar idea to downscale air quality models. We use the coherence function to illuminate fundamental limitations on some popular multivariate covariance constructions.

2.2. Implications for some existing models

The coherence function can be used as a tool to compare proposed multivariate models, as an indicator of the amount of flexibility of bivariate relationships at differing frequencies. For example, a common approach to specifying covariances is separability, setting $\mathbf{C}(\mathbf{h}) = \mathbf{R}C(\mathbf{h})$ where $C(\mathbf{h})$ is a univariate covariance function and $\mathbf{R}$ is a $p \times p$ positive definite matrix (Mardia and Goodall (1993); Helterbrand and Cressie (1994); Bhat, Haran and Goes (2010)). This approach has been empirically shown to be insufficiently flexible, and the following result contributes to the empirical results.

Proposition 1. *If* $\mathbf{C}(\mathbf{h}) = \mathbf{R}C(\mathbf{h})$ *where* $C : \mathbb{R}^d \to \mathbb{R}$ *is a covariance function and* $\mathbf{R}$ *is a* $p \times p$ *positive definite matrix with* (i, j)*th entry* r_{ij}*, then the squared coherence between the* i*th and* j*th process is constant, with* $\gamma_{ij}(\boldsymbol{\omega})^2 = (r_{ij}r_{ji})/(r_{ii}r_{jj})$.

This result can be contrasted against the analogous result for separable space-time processes (Fuentes (2006)), where process index can be viewed as indexing spatial coordinate and frequency being time-frequency rather than multidimensional. Such separation and constant coherence is the essential idea for the testing procedure developed by Fuentes (2006).

A more sophisticated method of generating multivariate covariance structures is to convolve univariate square integrable functions (Gaspari and Cohn (1999); Oliver (2003); Gaspari et al. (2006); Majumdar and Gelfand (2007)). In particular, if $c_i : \mathbb{R}^d \to \mathbb{R}$ are square integrable functions for $i = 1, \ldots, p$, then $C_{ij}(\mathbf{h}) = (c_i \star c_j)(\mathbf{h})$ is a valid matrix-covariance function where $\star$ denotes the convolution operator. This is sometimes known as covariance convolution (es-

pecially when the c_i are positive definite functions). This approach to model building is also overly-restrictive.

Proposition 2. *If c_1 and c_2 are square integrable functions on $\mathbb{R}^d$ and a matrix-valued covariance is defined via $C_{ij} = c_i \star c_j$ for $i, j = 1, 2$, where $\star$ denotes convolution, then $\gamma(\boldsymbol{\omega}) \equiv 1$ for all $\boldsymbol{\omega} \in \mathbb{R}^d$ such that the Fourier transforms of c_i and c_j are nonzero.*

Multivariate processes can sometimes be modeled as being related by local averaging. For example, the relationship between column integrated ozone observations and local ozone might plausibly be modeled as observations being locally averaged over the true underlying field (Cressie and Johannesson (2008)). Wind observations are often time averaged over moving windows to produce smoother and more stable observation series (Hering, Kazor and Kleiber (2015)). We characterize the coherence in such situations.

Proposition 3. *If $Z_1(\mathbf{s})$ is a weakly stationary stochastic process and $Z_2(\mathbf{s}) = \int_{\mathbb{R}^d} K(\mathbf{u} - \mathbf{s}) Z_1(\mathbf{u}) \mathrm{d}\mathbf{u}$ for some continuous square integrable kernel function $K : \mathbb{R}^d \to \mathbb{R}$ that is symmetric, then $\gamma(\boldsymbol{\omega}) \equiv 1$ for all $\boldsymbol{\omega} \in \mathbb{R}^d$ such that the Fourier transform of K is nonzero.*

According to Proposition 3, estimated coherences near unity over all frequency bands may be indicative of a linear or local averaged relationship between processes, and this result may serve as the theoretical basis for testing such a hypothesis. Fuentes (2006) developed a test for separability of space-time processes based on a similar notion. Proposition 3 is a multidimensional extension of exercise 8.6.16 of Brillinger (2001), and indeed follows from identical arguments as the time series result.

The kernel convolution method, introduced by Ver Hoef and Barry (1998) and Ver Hoef, Cressie and Barry (2004), originally involved representing a process as a moving average against a white noise process. In simple cases this yields the covariance convolution model. The basic form of the multivariate kernel convolution method involves integrating against correlated stochastic processes,

$$Z_k(\mathbf{s}) = \int_{\mathbb{R}^d} g_k(\mathbf{x} - \mathbf{s}) W_k(\mathbf{x}) \mathrm{d}\mathbf{x}, \tag{2.3}$$

where $W_1, \ldots, W_p$ are correlated stationary processes and $g_k : \mathbb{R}^d \to \mathbb{R}$ are square integrable symmetric kernel functions for $k = 1, \ldots, p$.

Proposition 4. *If $Z_1(\mathbf{s})$ and $Z_2(\mathbf{s})$ are constructed as in (2.3), with $W_1 = W_2$ almost everywhere, then $\gamma(\boldsymbol{\omega})$ is constant for all $\boldsymbol{\omega} \in \mathbb{R}^d$.*

For separable, covariance convolution, and simple cases of kernel convolution, the multivariate structures are restricted to constant coherence.

2.3. Phase and gain

Similar notions to frequency coherence can be motivated by examining spectral density matrices. If $(Z_1(\mathbf{s}), Z_2(\mathbf{s}))^{\mathrm{T}}$ is a stationary random vector with spectral density matrix $(f_{ij}(\boldsymbol{\omega}))_{i,j=1}^2$, take $A(\boldsymbol{\omega}) = f_{12}(\boldsymbol{\omega})/f_{11}(\boldsymbol{\omega})$, possibly complex-valued. The gain function $G(\boldsymbol{\omega}) = |A(\boldsymbol{\omega})|$ is sometimes referred to as the gain of $Z_2(\mathbf{s})$ on $Z_1(\mathbf{s})$, in time series (Brockwell and Davis (2009)). The phase function at frequency $\boldsymbol{\omega}$ is $\phi(\boldsymbol{\omega}) = \arg A(\boldsymbol{\omega})$. It satisfies $\phi(\boldsymbol{\omega}) \in (-\pi, \pi]$ and $\phi(-\boldsymbol{\omega}) = -\phi(\boldsymbol{\omega})$.

The interpretations of gain and phase are clearest when considering processes built by the relationship $Z_1(\mathbf{s}) = \alpha Z_2(\mathbf{s} - \mathbf{u})$ for some $\mathbf{u} \in \mathbb{R}^d$ and $\alpha \neq 0$. It is straightforward to show that the phase function is

$$\phi(\boldsymbol{\omega}) = \begin{cases} -\boldsymbol{\omega}^{\mathrm{T}}\mathbf{u} \, (\mathrm{mod} \ 2\pi), & \alpha > 0, \\ \pi - \boldsymbol{\omega}^{\mathrm{T}}\mathbf{u} \, (\mathrm{mod} \ 2\pi) & \alpha < 0. \end{cases}$$

Li and Zhang (2011) develop an approach to modeling this type of asymmetric cross-covariance behavior. This shows that their construction has a phase function that depends on the angle $\boldsymbol{\omega}^{\mathrm{T}}\mathbf{u}$, and may be used as an exploratory data approach or as the basis for a statistical test of whether a pair of spatial processes exhibit an asymmetric relationship. Li and Zhang (2011) use the empirical cross-correlation function to visually assess such asymmetric behavior. The gain function in this case is simply $G(\boldsymbol{\omega}) = |\alpha|$; all frequency components of Z_2 are exaggerated by an amount α for Z_1.

We consider some multivariate constructions that are particular to real-valued processes having real-valued spectral matrices. Any model with real-valued cross-spectral density has $\phi(\boldsymbol{\omega}) \equiv 0$, but a possibly non-trivial gain function. Thus, testing $\phi(\boldsymbol{\omega}) \equiv 0$ can be viewed as a test for a real-valued cross-spectral density.

2.4. Revisiting the multivariate Matérn

The multivariate Matérn is a model for matrix-valued covariance functions such that each marginal is a Matérn covariance function, and all cross-covariance functions are in the Matérn class (Gneiting, Kleiber and Schlather (2010); Apanaso-vich, Genton and Sun (2012)). Specifically, the multivariate Matérn has $C_{ii}(\mathbf{h}) = \sigma_i^2 \mathrm{M}(\mathbf{h} \, | \, \nu_i, a_i)$ for $i = j$, and $C_{ij}(\mathbf{h}) = \rho_{ij}\sigma_i\sigma_j \mathrm{M}(\mathbf{h} \, | \, \nu_{ij}, a_{ij})$ for $1 \leq i \neq j \leq p$.

Here, $\mathrm{M}(\mathbf{h} \,|\, \nu, a) = (2^{1-\nu}/\Gamma(\nu))(a\|\mathbf{h}\|)^{\nu}\mathrm{K}_{\nu}(a\|\mathbf{h}\|)$, where K_{ν} is a modified Bessel function of the second kind of order ν. Gneiting, Kleiber and Schlather (2010) and Apanasovich, Genton and Sun (2012) discuss restrictions on the parameters $\nu_i, a_i, \nu_{ij}, a_{ij}$, and ρ_{ij} that result in a valid model.

The Matérn class is popular due to the smoothness parameter $\nu > 0$ that continuously indexes smoothness of the sample paths of the process. In particular, sample paths are m times differentiable if and only if $\nu > m$, and there is an additional relationship between ν and the fractal dimension in that sample paths have dimension $\max(d, d+1-\nu)$ (Goff and Jordan (1988); Handcock and Stein (1993)). These interpretations and implications also hold in the multivariate case, where ν_i indexes the smoothness of the ith component process $Z_i(\mathbf{s})$. The parameters a_i act as range parameters, and control the rate of decay of spatial correlation away from the origin.

A standing issue with the multivariate Matérn is that the cross-covariance parameters, ν_{ij} and a_{ij} for $i \neq j$, do not have straightforward interpretations and implications analogous to the marginal smoothness and range interpretations, and indeed nowhere in the literature have these parameters been linked directly to process behavior. Coherence provides natural explanations for the effects of the cross-smoothness and range parameters.

The squared coherence function for a bivariate process with multivariate Matérn correlation is

$$\gamma(\boldsymbol{\omega})^2 = \rho^2 \frac{\Gamma(\nu_{12} + d/2)^2 \Gamma(\nu_1)\Gamma(\nu_2)}{\Gamma(\nu_1 + d/2)\Gamma(\nu_2 + d/2)\Gamma(\nu_{12})^2} \frac{a_{12}^{4\nu_{12}}}{a_1^{2\nu_1} a_2^{2\nu_2}}$$
$$\frac{(a_1^2 + \|\boldsymbol{\omega}\|^2)^{\nu_1 + d/2}(a_2^2 + \|\boldsymbol{\omega}\|^2)^{\nu_2 + d/2}}{(a_{12}^2 + \|\boldsymbol{\omega}\|^2)^{2\nu_{12} + d}}. \tag{2.4}$$

Cross-Smoothness ν_{12}

Restrictions on the cross-covariance parameters that result in a valid covariance model are characterized in Theorem 3 of Gneiting, Kleiber and Schlather (2010). In particular, if $\nu_{12} < (\nu_1 + \nu_2)/2$ then $\rho = 0$. If $\nu_{12} > (\nu_1 + \nu_2)/2$ then the restrictions on a_{12} and ρ are relatively complicated. The choice of $\nu_{12} = (\nu_1 + \nu_2)/2$ results in easy-to-check conditions on ρ, and removes a free parameter from the model. But, from (2.4), the squared coherence converges to the non-trivial value

$$\rho^2 \frac{\Gamma(\nu_{12} + d/2)^2 \Gamma(\nu_1)\Gamma(\nu_2)}{\Gamma(\nu_1 + d/2)\Gamma(\nu_2 + d/2)\Gamma(\nu_{12})^2} \frac{a_{12}^{4\nu_{12}}}{a_1^{2\nu_1} a_2^{2\nu_2}}$$

as $\|\boldsymbol{\omega}\| \to \infty$. Under $\nu_{12} = (\nu_1 + \nu_2)/2$, the two processes Z_1 and Z_2 are corre-
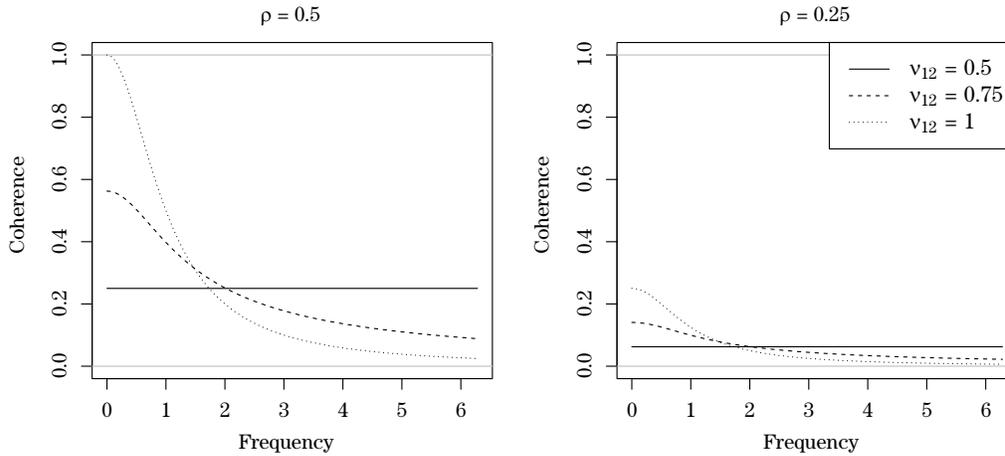
Figure 1. Squared coherence functions for various bivariate Matérns with $a_1 = a_{12} = a_2 = 1, \nu_1 = \nu_2 = 0.5$, varying ν_{12} and ρ.

lated at *arbitrarily high* frequencies, and may not be justifiable for some physical processes.

If $a_1 = a_2 = a_{12} = a$, all covariance and cross-covariance functions share a common range. In this case,

$$\gamma(\boldsymbol{\omega})^2 = \rho^2 \frac{\Gamma(\nu_{12} + d/2)^2 \Gamma(\nu_1)\Gamma(\nu_2)}{\Gamma(\nu_1 + d/2)\Gamma(\nu_2 + d/2)\Gamma(\nu_{12})^2} (a^2 + \|\boldsymbol{\omega}\|^2)^{\nu_1 + \nu_2 - 2\nu_{12}}. \qquad (2.5)$$

Figure 1 shows some coherence functions for common length scale parameters, varying ν_{12} and ρ. Clearly, ν_{12} controls the rate of decay at high frequencies, and indeed $\gamma(\boldsymbol{\omega})^2 \sim \|\boldsymbol{\omega}\|^{2\nu_1 + 2\nu_2 - 4\nu_{12}}$ at high frequencies when $\nu_{12} > (\nu_1 + \nu_2)/2$.

Cross-Range a_{12}

The cross-range parameter a_{12} controls the *frequency of greatest coherence*, which can either be at $\|\boldsymbol{\omega}\| = 0$, $\|\boldsymbol{\omega}\| \to \infty$, or for finite but nonzero frequencies. The frequency of greatest coherence can be derived by taking derivatives of the log coherence (2.4). In general, they occur at solutions to

$$\frac{\nu_1 + d/2}{a_1^2 + \|\boldsymbol{\omega}\|^2} + \frac{\nu_2 + d/2}{a_2^2 + \|\boldsymbol{\omega}\|^2} - \frac{2\nu_{12} + d}{a_{12}^2 + \|\boldsymbol{\omega}\|^2} = 0.$$

We discuss two special cases. Setting $a_1 = a_2 = a$ and $\nu_1 = \nu_2 = \nu$, we suppose $\nu_{12} > \nu$ (distinct a_1 and a_2 and ν_1 and ν_2 follow qualitatively similar behavior). Then the frequency of maximal coherence is

$$\|\boldsymbol{\omega}\|^2 = \frac{a_{12}^2(2\nu + d) - a^2(2\nu_{12} + d)}{2(\nu_{12} - \nu)}.$$

When the numerator is positive, the frequency of greatest coherence is linear
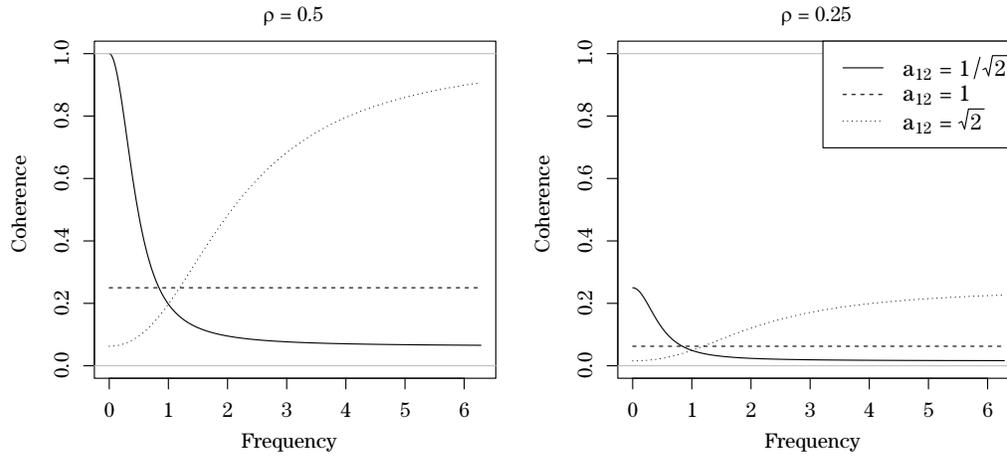
Figure 2. Squared coherence functions for various bivariate Matérns with $\nu_1 = \nu_{12} = \nu_2 = 1, a_1 = a_2 = 1$, varying a_{12} and ρ.

in a_{12}^2, and thus the effect of varying the cross-range is to move the location of maximal coherence. For $\nu_1 = \nu_2 = \nu_{12} = \nu$, and $a_1 = a_2 = a$, the coherence function is

$$\gamma(\boldsymbol{\omega})^2 = \rho^2 \left(\frac{a_{12}^2}{a^2}\right)^{2\nu} \left(\frac{a^2 + \|\boldsymbol{\omega}\|^2}{a_{12}^2 + \|\boldsymbol{\omega}\|^2}\right)^{2\nu+d}. \tag{2.6}$$

If $a_{12} < a$, maximal coherence is at $\|\boldsymbol{\omega}\| = 0$, whereas if $a_{12} > a$, the maximal coherence is at $\|\boldsymbol{\omega}\| \to \infty$. Figure 2 illustrates these scenarios.

The multivariate Matérn is flexible in that it can put greatest coherence at zero or infinity, and can allow for increasing, decreasing or constant coherence. Not all choices imply monotonic coherence, though; for example, various choices of parameters can imply a 'bump' in coherence that is between zero and infinity, for example see Figure 3.

The so-called parsimonious Matérn model arises from imposing common range parameters and taking $\nu_{12} = (\nu_1 + \nu_2)/2$ (Gneiting, Kleiber and Schlather (2010)). This model has been empirically shown to produce inferior model fits to datasets as compared to more general versions of the multivariate Matérn as well as other multivariate classes (Apanasovich, Genton and Sun (2012); Genton and Kleiber (2015)). The coherence function for a bivariate parsimonious Matérn model is constant, $\gamma(\boldsymbol{\omega}) = \rho$, which suggests an inflexible model for the spectral behavior of spatial processes.

2.5. The linear model of coregionalization

The linear model of coregionalization (LMC) is built by decomposing a mul-
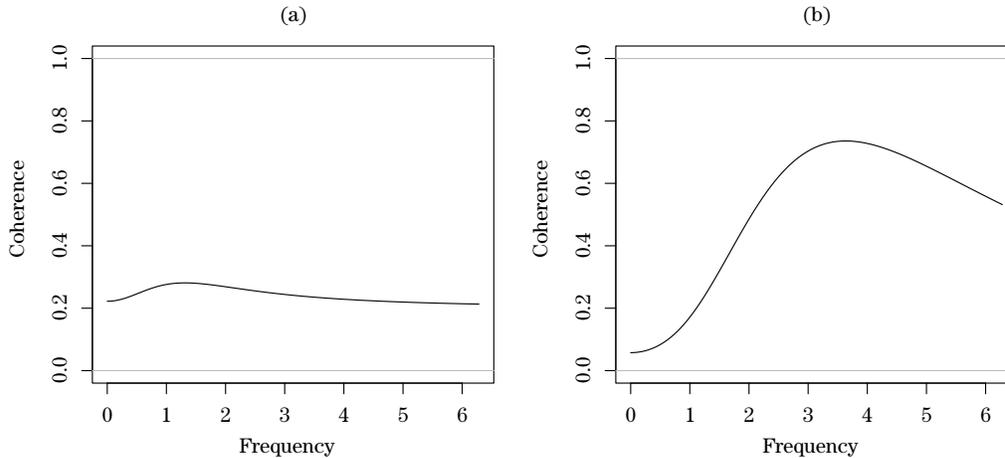
Figure 3. Squared coherence functions for two bivariate Matérns with (a) $\nu_1 = 1, \nu_{12} = 1.5, \nu_2 = 2, a_1 = 1, a_2 = 2, a_{12} = 1.5$; (b) $\nu_1 = 1, \nu_{12} = 1.5, \nu_2 = 1, a_1 = 1, a_2 = 2, a_{12} = 2.5$; $\rho = 0.5$ in both cases.

tivariate process as linear combinations of uncorrelated, univariate processes (Goulard and Voltz (1992); Royle and Berliner (1999); Wackernagel (2003); Schmidt and Gelfand (2003)). In particular, we entertain version

$$\mathbf{Z}(\mathbf{s}) = \left(\begin{array}{c} Z_1(\mathbf{s}) \\ Z_2(\mathbf{s}) \end{array} \right) = \left(\begin{array}{cc} b_{11} & b_{12} \\ b_{21} & b_{22} \end{array} \right) \left(\begin{array}{c} W_1(\mathbf{s}) \\ W_2(\mathbf{s}) \end{array} \right) = \mathbf{B}\mathbf{W}(\mathbf{s}). \qquad (2.7)$$

The matrix $\mathbf{B}$ is known as the coregionalization matrix, and controls the strength of dependencies on the latent uncorrelated processes $\mathbf{W}$. Here, suppose $W_1(\mathbf{s})$ and $W_2(\mathbf{s})$ are uncorrelated processes with spectral densities $f_1(\boldsymbol{\omega})$ and $f_2(\boldsymbol{\omega})$, respectively.

Given the number of parameters in the LMC, it is often useful to impose restrictions on the coregionalization matrix $\mathbf{B}$, such as setting $b_{11} = b_{22} = 1$ (Berrocal, Gelfand and Holland (2010)).

Lemma 1. *In the linear model of coregionalization* (2.7), *if $b_{11} = b_{22} = 1$ then the coherence function between $Z_1(\mathbf{s})$ and $Z_2(\mathbf{s})$ is unity if and only if $b_{12}b_{21} = 1$.*

Thus, under the LMC, two processes are exactly coherent when they differ only by a scalar multiplier.

How does the coherence depend on b_{12} and b_{21} at high frequencies? Without loss of generality, suppose W_1 has a spectral density that decays as $\|\boldsymbol{\omega}\|^{-2\nu_1-d}$ as $\|\boldsymbol{\omega}\| \to \infty$, and W_2 has a spectral density that decays as $\|\boldsymbol{\omega}\|^{-2\nu_2-d}$ as $\|\boldsymbol{\omega}\| \to \infty$ where $\nu_1 < \nu_2$ (we can interpret W_1 as the rough process and W_2 as the smooth

process). It is straightforward to show that the coherence between Z_1 and Z_2 converges to a nonzero constant as $\|\boldsymbol{\omega}\| \to \infty$ if $b_{21} \neq 0$, whereas it converges to zero if $b_{12} = 0$. This suggests that the coregionalization matrix should be structured to include off-diagonal zeros, so that Z_1 and Z_2 do not both share components of W_1 and W_2, unless there is a physically justifiable argument for nontrivial coherence at arbitrarily high frequencies.

Under the assumption $b_{11} = b_{22} = 1$, we have the gain function of $Z_2(\mathbf{s})$ on $Z_1(\mathbf{s})$ is

$$G(\boldsymbol{\omega}) = \frac{b_{21}f_1(\boldsymbol{\omega}) + b_{12}f_2(\boldsymbol{\omega})}{f_1(\boldsymbol{\omega}) + b_{12}^2 f_2(\boldsymbol{\omega})}.$$

If, as is common in using the LMC, we set $b_{12} = 0$, then there is constant gain at all frequencies by the amount of coregionalization, b_{21}. The complementary case where $b_{21} = 0$ yields the gain

$$G(\boldsymbol{\omega}) = \frac{b_{12}f_2(\boldsymbol{\omega})}{f_1(\boldsymbol{\omega}) + b_{12}^2 f_2(\boldsymbol{\omega})},$$

the relative contribution of component $b_{12}W_2(\mathbf{s})$ to the combined spectrum of $Z_1(\mathbf{s})$. If the latent processes have real-valued spectral densities, the phase function is exactly zero at all frequencies.

3. Estimation of Spectra

Suppose $\mathbf{Z}(\mathbf{s})$ is a p-variate process that has been observed at a regular grid of points $\{\mathbf{s}_i\}_{i=1}^N$, of marginal dimensions $n_i, i = 1, \ldots, d$ where $N = \prod_{i=1}^d n_i$. If grid spacing in the ith dimension is δ_i, take $\delta = \prod_{i=1}^d \delta_i$. Then the spatial periodogram matrix is defined as $\mathbf{I}(\boldsymbol{\omega}) = (I_{k\ell}(\boldsymbol{\omega}))_{k,\ell=1}^p$ where

$$I_{k\ell}(\boldsymbol{\omega}) = \frac{\delta}{(2\pi)^d N} \left(\sum_{k=1}^N Z_k(\mathbf{s}_k) \exp(-i\mathbf{s}_k^{\mathrm{T}}\boldsymbol{\omega}) \right) \overline{\left(\sum_{k=1}^N Z_\ell(\mathbf{s}_k) \exp(-i\mathbf{s}_k^{\mathrm{T}}\boldsymbol{\omega}) \right)}, \quad (3.1)$$

and is available at Fourier frequencies $\boldsymbol{\omega} = 2\pi\mathbf{f}$ where $\mathbf{f} = (f_1/(\delta_1 n_1), \ldots, f_d/(\delta_d n_d))^{\mathrm{T}}$ for $f_i \in \{-\lfloor (n_i - 1)/2 \rfloor, \ldots, n_i - \lfloor n_i/2 \rfloor\}$. Note that $I_{k\ell}(\boldsymbol{\omega}) = \overline{I_{\ell k}(\boldsymbol{\omega})}$.

It is natural to consider asymptotics for time series as time $t \to \infty$, resulting in effectively uncorrelated blocks of a process; in the spatial realm there are two competing asymptotic frameworks. Increasing domain asymptotics has samples taken on an ever-increasing domain in all axial directions, and typically asymptotic results here echo those in time series. Infill asymptotics has the domain boundary fixed and points sampled at an ever finer resolution within the domain (Zhang and Zimmerman (2005)). The large sample properties of the periodogram

(3.1) change accordingly.

Using infill asymptotics, Lim and Stein (2008) show that the raw multivariate periodogram can exhibit bias at low frequencies, and suggest prewhitening the process to overcome this inadequacy (in the univariate case Stein (1995) gives a simulated example where the bias is quite substantial). However, under a mixture of infill and increasing domain asymptotics, Fuentes (2002) showed (for univariate processes) the analogous result to the time series case that the periodogram is asymptotically unbiased and is uncorrelated at differing Fourier frequencies. Additionally, in this latter case it is not a consistent estimator, but must be smoothed to gain consistency.

Under certain assumptions the nonparametric periodogram (3.1) is asymptotically unbiased, and generates asymptotically uncorrelated random variables between distinct Fourier frequencies. In this, use the same assumptions as Fuentes (2002), generalized to the multivariate setting.

A1 The true spectral densities $f_{k\ell}(\boldsymbol{\omega})$ decay as $\|\boldsymbol{\omega}\|^{-\tau}, \tau > 2$ as $\|\boldsymbol{\omega}\| \to \infty$, $\boldsymbol{\omega} \in \mathbb{R}^2$.

A2 The marginal and cross-covariances satisfy $\int \|\mathbf{h}\| |C_{k\ell}(\mathbf{h})| \mathrm{d}\mathbf{h} < \infty, \mathbf{h} \in \mathbb{R}^2$.

A3 $\delta_i \to 0$, $n_i \to \infty$ and $\delta_i n_i \to \infty$ for all $i, j = 1, 2$ such that $n_i/n_j \to \lambda_{ij} > 0$.

Theorem 3. *Under* A1-A3, *we have* $\mathbb{E}I_{k\ell}(\boldsymbol{\omega}) \to f_{k\ell}(\boldsymbol{\omega})$, $\mathrm{Var}I_{k\ell}(\boldsymbol{\omega}) \to f_{k\ell}(\boldsymbol{\omega})^2$ *and* $\mathrm{Cov}(I_{k\ell}(\boldsymbol{\omega}_1), I_{k\ell}(\boldsymbol{\omega}_2)) \to 0$ *for* $\boldsymbol{\omega}_1 \neq \boldsymbol{\omega}_2$.

The proof for Theorem 3 follows directly from Fuentes (2002) and is not included here.

According to Theorem 3, the matrix-valued periodogram is not an asymptotically consistent estimator. To produce a consistent estimator of the spectral density at a particular frequency $\boldsymbol{\omega}_0$, in practice we locally smooth adjacent periodogram values and appeal to Theorem 3. In particular, the smoothed matrix-valued periodogram is

$$\tilde{I}_{k\ell}(\boldsymbol{\omega}_0) = \int_{\mathbb{R}^d} K_\lambda(\boldsymbol{\omega} - \boldsymbol{\omega}_0) I_{k\ell}(\boldsymbol{\omega}) \mathrm{d}F_n(\boldsymbol{\omega}), \tag{3.2}$$

where $F_n(\boldsymbol{\omega})$ is the empirical cumulative distribution function of Fourier frequencies $\{\boldsymbol{\omega}_i\}_{i=1}^N$. Here, K_λ is some kernel function with bandwidth λ, where, as we have it written, the same kernel is applied to each process. Naturally, different kernels may be used for different processes if the scientific context calls for such an approach.

Table 1. Empirical bias and root mean squared error (RMSE) between empirical and theoretical coherence for various bivariate Matérn correlation models for sample sizes of n^2 for $n = 50, 100, 200$. Parenthetical values are standard errors based on 100 independent simulations.

	$a = 1, \nu_1 = 0.5, \nu_2 = 0.5, \nu_{12} = 0.5, \rho = 0.5$			$a = 1, \nu_1 = 0.5, \nu_2 = 0.5, \nu_{12} = 1.5, \rho = 0.3$		
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
Bias	0.002 (0.030)	0.001 (0.016)	0.001 (0.009)	−0.041 (0.022)	−0.038 (0.012)	−0.032 (0.006)
RMSE	0.04 (0.02)	0.03 (0.01)	0.02 (0.01)	0.13 (0.02)	0.09 (0.01)	0.07 (0.01)
	$a = 1, \nu_1 = 0.5, \nu_2 = 0.5, \nu_{12} = 2.5, \rho = 0.15$			$a = 5, \nu_1 = 0.5, \nu_2 = 0.5, \nu_{12} = 0.5, \rho = 0.5$		
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
Bias	−0.027 (0.015)	−0.026 (0.009)	−0.019 (0.005)	0.001 (0.029)	0.001 (0.015)	0.001 (0.01)
RMSE	0.11 (0.01)	0.09 (0.01)	0.06 (0.01)	0.04 (0.01)	0.03 (0.01)	0.02 (0.01)
	$a = 0.2, \nu_1 = 0.5, \nu_2 = 0.5, \nu_{12} = 0.5, \rho = 0.5$			$a = 1, \nu_1 = 0.5, \nu_2 = 1, \nu_{12} = 1.5, \rho = 0.4$		
	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
Bias	0.004 (0.053)	0.001 (0.028)	0.003 (0.013)	−0.017 (0.024)	−0.021 (0.013)	−0.020 (0.007)
RMSE	0.07 (0.03)	0.05 (0.02)	0.02 (0.01)	0.09 (0.02)	0.07 (0.01)	0.04 (0.01)

We cannot directly use the nonparametric periodogram fraction to estimate the coherence as $I_{k\ell}(\boldsymbol{\omega})I_{\ell k}(\boldsymbol{\omega}) = I_{kk}(\boldsymbol{\omega})I_{\ell\ell}(\boldsymbol{\omega})$ at all Fourier frequencies. Thus, we estimate the coherence functions by using the smoothed periodograms,

$$\hat{\gamma}_{k\ell}(\boldsymbol{\omega})^2 = \frac{|\tilde{I}_{k\ell}(\boldsymbol{\omega})|^2}{\tilde{I}_{kk}(\boldsymbol{\omega})\tilde{I}_{\ell\ell}(\boldsymbol{\omega})}$$

for $k, \ell = 1, \ldots, p$.

We close this section with a small simulation study to build some intuition for the effect of sample size and model parameters on estimating the coherence function. We considered sample sizes of n^2 for $n = 50, 100, 200$, and various multivariate Matérn covariance models. For each sample size, we generated 100 realizations of a bivariate Gaussian process on a grid of coordinates $\{(i, j)\}_{i,j=1}^n$ and estimated the coherence function using a two-dimensional Daniell kernel with a bandwidth of $\sqrt{n}$. For each simulation we calculated the average bias and root mean squared error (RMSE) over all Fourier frequencies, and then averaged over simulations. Table 1 reports these values of bias and RMSE, along with standard errors.

4. Illustrations

We examine two datasets from the atmospheric sciences, gridded reforecasts and reanalyses of sea level pressure and geopotential heights over the equatorial region. Reforecast data are produced retrospectively from a fixed version of a numerical weather prediction model, in this case the 2nd generation National Oceanic and Atmospheric Administration's (NOAA) Global Ensemble Forecast System Reforecast (Hamill et al. (2013)). Forecasts are generated at 3 hour incre-

ments from 0 to 192 hours, with the 0h forecast being a reanalysis, an estimate of the current state of the atmosphere. For our data, the control initial conditions were produced using a hybrid ensemble Kalman filter-variational analysis system (Hamill et al. (2011)). Reanalysis data are particularly useful and important as many physical processes (such as geopotential heights) are extremely difficult to observe directly (e.g., geopotential heights are typically measured by atmospheric soundings at a handful of locations worldwide).

4.1. Sea level pressure

The first dataset we considered was a set of reforecast sea level pressures (SLP) over the equatorial region. Sea level pressures in this region are approximately stationary, and we compare forecast horizons in 24 hour increments from 0h to 192h (8 days out). The data consist of gridded reforecasts from the first 90 days of 2014 at $1°$ increments over 360 longitude and 47 latitude bands between $-23°$ to $23°$, defining the equatorial region.

One approach to examining the quality of forecasts is the coherence between the forecast with the corresponding reanalysis. For example, we might compare the 24h forecast of SLP generated on January 1, 2014 to the 0h reanalysis generated on January 2, 2014. It is well known that forecast skill decays with horizon, and we expect the short-term forecasts to share higher coherence with the reanalyses than the long-term forecasts.

We began by standardizing each analysis and forecast horizon grid cell by subtracting the temporal average and dividing by empirical standard deviation to produce forecast anomalies. Denote these anomalies by $Z_k(\mathbf{s}, d)$ for forecast horizons $k = 0, 1, 2, \ldots, 8$ corresponding to forecast horizons $0, 24, \ldots, 192$ hours, spatial locations $\mathbf{s} \in \mathcal{D} \subset \mathbb{R}^2$ in the equatorial region $\mathcal{D}$ on days $d = 1, \ldots, 90$.

Each day's marginal process empirical periodogram, (3.1), was calculated for all forecast horizons k, yielding $\{I_{kk}(\boldsymbol{\omega}, d)\}$. The smoothed periodogram was a convolution with a simple low-pass filter, a matrix of zeros with a 3×3 constant block of $1/9$. Interest focused on comparing various forecast horizons with the reanalysis at $k = 0$, so we calculated empirical cross-periodograms $\{I_{0k}(\boldsymbol{\omega}, d)\}$ for all available days d allowing for forecast validation (e.g., the $k = 1$, 24h horizon, has 89 available days, $d = 2, \ldots, 90$). The cross-periodograms were smoothed using the same low-pass filter as the marginals. With $\tilde{I}_{k\ell}(\boldsymbol{\omega}, d)$ as the smoothed cross-periodograms, we estimated the squared coherence function as

$$\hat{\gamma}_{0k}(\boldsymbol{\omega})^2 = \frac{1}{90 - k} \sum_{d=1+k}^{90} \frac{|\tilde{I}_{0k}(\boldsymbol{\omega}, d)|^2}{\tilde{I}_{00}(\boldsymbol{\omega}, d)\tilde{I}_{kk}(\boldsymbol{\omega}, d - k)},$$
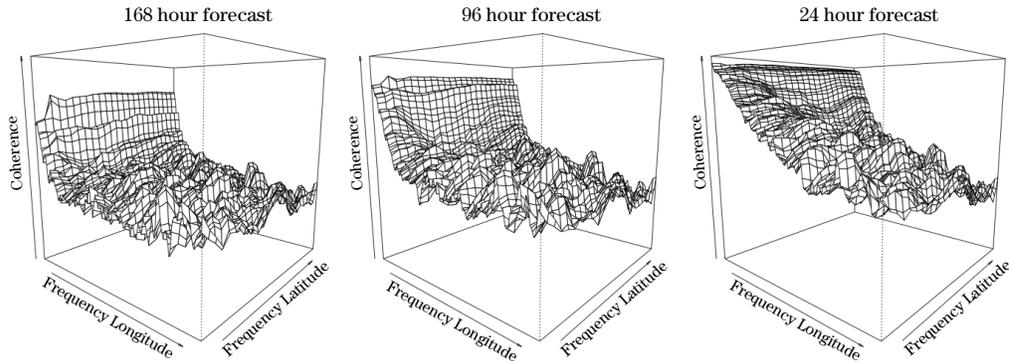
Figure 4. Estimated absolute coherence functions for the GEFS sea level pressure re-forecast data, comparing the 168, 96 and 24 hour forecast horizons with the zero hour analysis. The vertical axis spans $[0, 1]$.

for $k = 1, \ldots, 8$, that is, the average over all available daily smoothed cross-periodograms.

Figure 4 shows estimated absolute coherence functions for horizons 168, 92, and 24h with the 0h analysis. Even at long lead times there is substantial coherence, which increases by a substantial margin at very low longitudinal frequencies. For any given longitude frequency band, the coherence appears to be relatively constant across latitudes, which is sensible given that there appears to be greater variability in the equatorial direction than in the north-south direction for sea level pressure in this region. As the forecast horizon decreases the coherence begins building between low-to-mid frequency bands in the latitudinal direction, suggesting that the statistical characteristics of short term forecasts are more similar to observed sea level pressure than are the longer term forecasts. At the highest frequencies there is not a substantial improvement in forecast skill, bordering on no improvement, which suggests that small scale events are difficult to forecast even at one day out.

To gain a better understanding of where the forecasts are improving as the horizon decreases toward zero, we considered pairwise differences of coherence functions. Figure 5 displays maps of pointwise difference between adjacent forecast horizons. For example, Panel (a) shows the improvement in coherence in moving from a 48h to a 24h forecast horizon. Apparently, the greatest improvements in forecast ability begin to occur at approximately 2-3 days out. Moreover, as the validation time nears, the forecasts improve at higher frequencies. Such an exploratory analysis would be difficult to accomplish with standard covariance techniques.
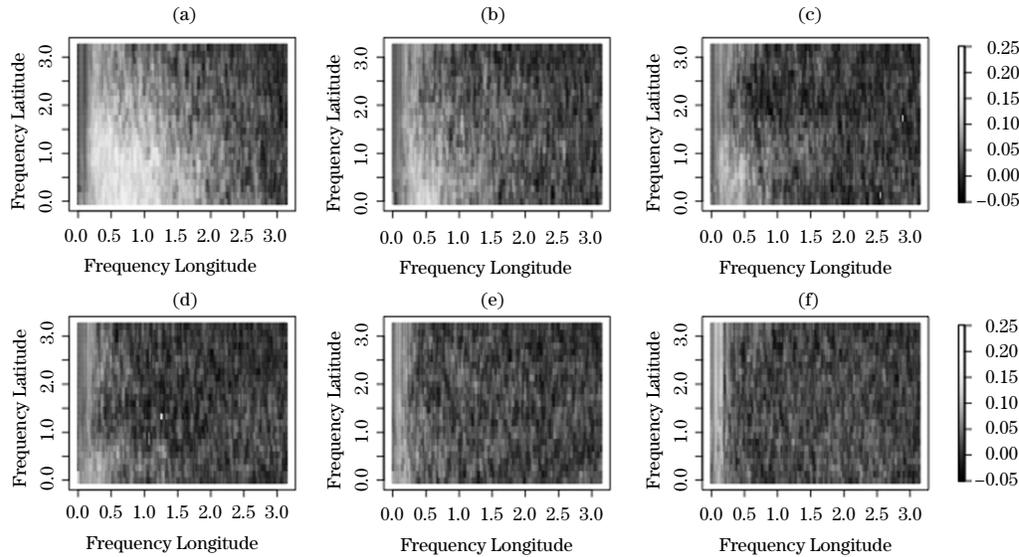
Figure 5. Incremental differences in coherence for the GEFS sea level pressure reforecast data. Panels show the improvement (as a difference) in coherence for (a) 48 to 24h, (b) 72 to 48h, (c) 96 to 72h, (d) 120 to 96h, (e) 144 to 120h, and (f) 168 to 144h.

4.2. Geopotential height

Our second example is on the same spatial domain, but the data are geopotential heights. Geopotential height is the height (in meters) above sea level at which the atmospheric pressure is a certain level. In the atmospheric sciences, it is common to examine geopotential heights as indicators of climatic regimes; for instance, Knapp and Yin (1996) discuss the relationship between heights and temperature anomalies over a portion of the United States.

Three common geopotential height maps are the 850hPa, 500hPa, and 300hPa maps. The first, 850hPa, approximately defines the planetary boundary layer, the lowest level of the atmosphere that interacts with the surface of the Earth (note 1000hPa is approximately sea level). The 300hPa level is at the core of the jet stream, while the 500hPa approximately divides the atmosphere in half, and whose anomalies are used in part to assess climatological temperature variations. The vertical structure of geopotential heights is a focus of some interest within atmospheric sciences (Blackmon et al. (1979)).

We examined geopotential height reanalysis anomalies $Z_k(\mathbf{s}, d)$ for $k = 1, 2, 3$, representing the 850hPa, 500hPa, and 300hPa pressure levels on days $d = 1, \ldots$, 181. The anomalies are differences between the reanalysis height and a time-varying Nadaraya-Watson kernel smoothed estimate of the mean with a band-
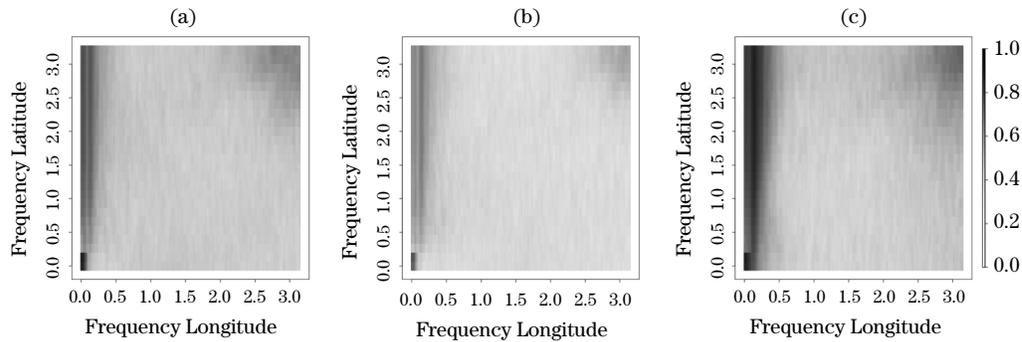
Figure 6. Estimated absolute coherence functions for the GEFS geopotential height data between (a) 850hPa and 500hPa, (b) 850hPa and 300hPa and, (c) 500hPa and 300hPa.

width of 5 days. Experiments suggested that results are qualitatively robust against choices of the bandwidth and smoothing kernel.

We smoothed the marginal process periodogram (3.1) using a low-pass filter, and calculated smoothed empirical cross-periodograms yielding $\{\tilde{I}_{ij}(\boldsymbol{\omega}, d)\}_{i,j=1}^{3}$. Then the squared coherence was estimated as the arithmetic average of each day's empirical squared coherence estimate,

$$\hat{\gamma}_{ij}(\boldsymbol{\omega})^2 = \frac{1}{181} \sum_{d=1}^{181} \frac{|\tilde{I}_{ij}(\boldsymbol{\omega}, d)|^2}{\tilde{I}_{ii}(\boldsymbol{\omega}, d) \tilde{I}_{jj}(\boldsymbol{\omega}, d)}.$$

Figure 6 shows the three estimated pairwise absolute coherence functions. There is high coherence between the lower pressure levels at low frequencies, and some evidence of moderate coherence between all levels at low frequencies. We also note some strikingly different behavior than for the sea level pressure example. First, there is an apparent ridge in coherence at low frequencies (approximately $2\pi 9/360$) which may be indicative of equatorial planetary waves (Wang and Xie (1996); Xie and Wang (1996); Kiladis et al. (2009)). Planetary waves can play crucial roles in the formation of tropical cyclones (Molinari, Lombardo and Vollaro (2007)). Additionally, there is high coherence at the greatest Fourier frequencies simultaneously in both dimensions for all coherence functions (capping out at approximately 0.62, 0.54, and 0.70). This is evidence of a nonseparable relationship in the frequency domain, and we are unaware of any current multivariate models that can adequately capture such behavior. One possible explanation for this high coherence at high frequencies is artifacts in the data assimilation scheme, in particular aberrant observational data leading to unusually large anomalies in geopotential height. Indeed, variables such as sea level pressure are well constrained by a wealth of observational data, while geopotential
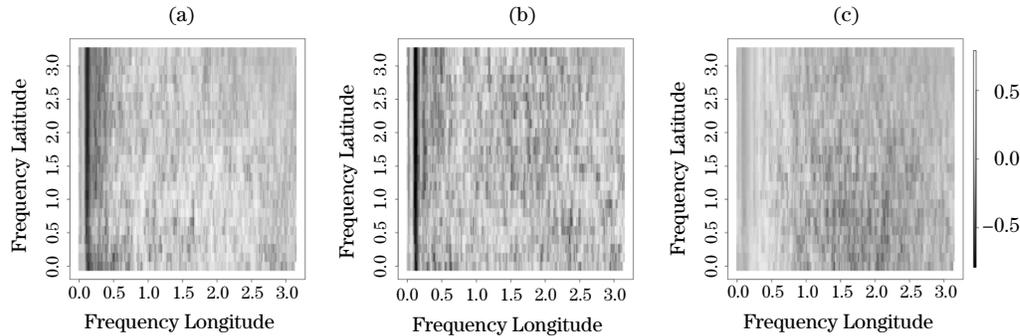
Figure 7. Estimated phase functions for the GEFS geopotential height data between (a) 850hPa and 500hPa, (b) 850hPa and 300hPa and (c) 500hPa and 300hPa.

heights are less constrained, usually being observed by sparsely released weather balloons.

Figure 7 shows pairwise plots for each pair of geopotential height anomalies. In particular there seems to be evidence of nontrivial phase at a low frequency band between the 850hPa and both lower pressure heights. These frequencies indicate longitudinal wavelengths of approximately 4,000–6,000 km, which is a typical wavelength for planetary equatorial waves, or Rossby waves (Wang and Xie (1996); Xie and Wang (1996); Kiladis et al. (2009)). Moreover, Panel (a) suggests there may be additional areas of positive phase at mid longitudinal and low-to-mid latitudinal frequencies, although we currently have no physical intuition for what these might represent. Most extant multivariate models utilize real-valued cross-spectral densities, and thus are insufficiently flexible to capture this type of phase behavior at specific spectra.

5. Discussion

The notions of coherence, phase and gain are common in the time series literature, but have not received much attention for multivariate spatial processes. The coherence between two variables can be interpreted as a measure of linear relationship at particular frequency bands, resulting in a complementary framework for comparing processes than the usual cross-covariance function. Phase and gain also yield straightforward interpretations as a physical space-shift and relative amplitude of frequency dependence when comparing two processes. We developed these ideas for stationary processes, and future research may be directed toward the analogous cases for nonstationary processes, perhaps extending the work of Fuentes (2002).

Coherence, phase, and gain can be estimated using smoothed cross-periodograms and, in our examples, we showed that as exploratory tools, these functions can be very useful in detecting structure that may not be readily captured using extant multivariate models. We additionally illustrated that the coherence function gives a natural interpretation to the multivariate Matérn cross-covariance parameters that have otherwise been uninterpretable.

A number of future research directions can be considered, including the adaptation of coherence to multivariate space-time processes. Development of frequency-nonseparable models is also of interest. Naturally, estimation at irregularly-sited locations is of particular interest for many spatial datasets, and could be an additional route for future research.

Supplementary Materials

Proofs for the main theorems and propositions are in the supplemental material.

Acknowledgment

References

Amblard, P., Coeurjolly, J., Lavancier, F. and Philippe, A. (2012). Basic properties of the multivariate fractional Brownian motion. *Bulletin de la Société Mathématique de France* **28**, 65–87.

Apanasovich, T. V., Genton, M. G. and Sun, Y. (2012). A valid Matérn class of cross-covariance functions for multivariate random fields with any number of components. *Journal of the American Statistical Association* **107**, 180–193.

Berrocal, V. J., Gelfand, A. E. and Holland, D. M. (2010). A bivariate space-time downscaler under space and time misalignment. *Annals of Applied Statistics* **4**, 1942–1975.

Bevilacqua, M., Hering, A. S. and Porcu, E. (2015). On the flexibility of multivariate covariance models: Comment on the paper by Genton and Kleiber. *Statistical Science* **30**, 167–169.

Bhat, K. S., Haran, M. and Goes, M. (2010). Computer model calibration with multivariate spatial output: a case study in climate parameter learning. In *Frontiers of Statistical Decision Making and Bayesian Analysis* (eds. M. H. Chen, P. Müller, D. Sun, K. Ye, and D. K. Dey) 401–408. Springer-Verlag, New York.

Blackmon, M. L., Madden, R. A., Wallace, J. M. and Gutzler, D. S. (1979). Geographical

variations in the vertical structure of geopotential height fluctuations. *Journal of the Atmospheric Sciences* **36**, 2450–2466.

Brillinger, D. R. (2001). *Time Series*, Society for Industrial and Applied Mathematics, Philadelphia, PA.

Brockwell, P. J. and Davis, R. A. (2009). *Time Series: Theory and Methods.* Springer Science & Business Media.

Carter, W. H. and Wolf, E. (1977). Coherence and radiometry with quasihomogeneous planar sources. *Journal of the Optical Society of America* **67**, 785–796.

Cramér, H. (1940). On the theory of stationary random processes. *Annals of Mathematics* **41**, 215–230.

Cressie, N. and Johannesson, G. (2008). Fixed rank kriging for very large spatial data sets. *Journal of the Royal Statistical Society* **70**, 209–226.

Fante, R. L. (1974). Mutual coherence function and frequency spectrum of a laser beam propagating through atmospheric turbulence. *Journal of the Optical Society of America* **64**, 592–598.

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

Fuentes, M. (2006). Testing for separability of spatial-temporal covariance functions. *Journal of Statistical Planning and Inference* **136**, 447–466.

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.

Gaspari, G., Cohn, S. E., Guo, J. and Pawson, S. (2006). Construction and application of covariance functions with variable length fields. *Quarterly Journal of the Royal Meteorological Society* **132**, 1815–1838.

Genton, M. G. and Kleiber, W. (2015). Cross-covariance functions for multivariate geostatistics. *Statistical Science* **30**, 147–163.

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.

Goff, J. A. and Jordan, T. H. (1988). Stochastic modeling of seafloor morphology: Inversion of sea beam data for second-order statistics. *Journal of Geophysical Research* **93**, 13589–13608.

Goulard, M. and Voltz, M. (1992). Linear coregionalization model: Tools for estimation and choice of cross-variogram matrix. *Mathematical Geology* **24**, 269–282.

Guinness, J., Fuentes, M., Hesterberg, D. and Polizzotto, M. (2014). Multivariate spatial modeling of conditional dependence in microscale soil elemental composition data. *Spatial Statistics* **9**, 93–108.

Hamill, T. M., Whitaker, J. S., Kleist, D. T., Fiorino, M. and Benjamin, S. G. (2011). Predictions of 2010's tropical cyclones using the GFS and ensemble-based data assimilation methods. *Monthly Weather Review* **139**, 3243–3247.

Hamill, T. M., Bates, G. T., Whitaker, J. S., Murray, D. R., Fiorino, M., Galarneau Jr., T. J., Zhu, Y. and Lapenta, W. (2013). NOAA's second-generation global medium-range ensemble reforecast dataset. *Bulletin of the American Meteorological Society* **94**, 1553–1565.

Handcock, M. S. and Stein, M. L. (1993). A Bayesian analysis of kriging. *Technometrics* **35**, 403–410.

Helterbrand, J. D. and Cressie, N. (1994). Universal co-kriging under intrinsic coregionalization. *Mathematical Geology* **26**, 205–226.

Hering, A. S., Kazor, K. and Kleiber, W. (2015). A Markov-switching vector autoregressive stochastic wind generator for multiple spatial and temporal scales. *Resources* **4**, 70–92.

Katz, E. J. and Briscoe, M. G. (1979). Vertical coherence of the internal wave field from towed sensors. *Journal of Physical Oceanography* **9**, 518–530.

Kiladis, G. N., Wheeler, M. C., Haertel, P. T., Straub, K. H. and Roundy, P. E. (2009). Convectively coupled equatorial waves. *Reviews of Geophysics* **47**.

Kleiber, W. and Nychka, D. W. (2015). Equivalent kriging. *Spatial Statistics* **12**, 31–49.

Knapp, P. A. and Yin, Z. (1996). Relationships between geopotential heights and temperature in the south-eastern US during wintertime warming and cooling periods. *International Journal of Climatology* **16**, 195–211.

Kneer, F. J., Mattig, W., Nesis, A. and Werner, W. (1980). Coherence analysis of granular intensity. *Solar Physics* **68**, 31–39.

Koopmans, L. H. (1964). On the coefficient of coherence for weakly stationary stochastic processes. *The Annals of Mathematical Statistics* pp. 532–549.

Li, B. and Zhang, H. (2011). An approach to modeling asymmetric multivariate spatial covariance structures. *Journal of Multivariate Analysis* **102**, 1445–1453.

Lim, C. Y. and Stein, M. (2008). Properties of spatial cross-periodograms using fixed-domain asymptotics. *Journal of Multivariate Analysis* **99**, 1962–1984.

Mack, H. and Flinn, E. A. (1971). Analysis of the spatial coherence of short-period acoustic-gravity waves in the atmosphere. *Geophysical Journal of the Royal Astronomical Society* **26**, 255–269.

Majumdar, A. and Gelfand, A. E. (2007). Multivariate spatial modeling for geostatistical data using convolved covariance functions. *Mathematical Geology* **39**, 225–245.

Mandel, L. and Wolf, E. (1976). Spectral coherence and the concept of cross-spectral purity. *Journal of the Optical Society of America* **66**, 529–535.

Mardia, K. and Goodall, C. (1993). Spatial-temporal analysis of multivariate environmental monitoring data. In *Multivariate Environmental Statistics*, (eds. G. P. Patil and C. R. Rao), 347–386, North Holland, Amsterdam.

Molinari, J., Lombardo, K. and Vollaro, D. (2007). Tropical cyclogenesis within an equatorial Rossby wave packet. *Journal of the Atmospheric Sciences* **64**, 1301–1317.

Oliver, D. S. (2003). Gaussian cosimulation: Modelling of the cross-covariance. *Mathematical Geology* **35**, 681–698.

Reich, B. J., Chang, H. H. and Foley, K. M. (2014). A spectral method for spatial downscaling. *Biometrics* **70**, 932–942.

Royle, A. and Berliner, L. M. (1999). A hierarchical approach to multivariate spatial modeling and prediction. *Journal of Agricultural, Biological and Environmental Statistics* **4**, 1–28.

Schmidt, A. M. and Gelfand, A. E. (2003). A Bayesian coregionalization approach for multivariate pollutant data. *Journal of Geophysical Research – Atmospheres* **108**.

Schreiner, A. E. and Dorman, L. M. (1990). Coherence lengths of seafloor noise: Effect of ocean bottom structure. *Journal of the Acoustic Society of America* **88**, 1503–1514.

Shumway, R. H. and Stoffer, D. S. (2011). *Time Series Analysis and Its Applications*. Springer: New York.

Stein, M. L. (1995). Fixed-domain asymptotics for spatial periodograms. *Journal of the American Statistical Association* **90**, 1277–1288.

Stein, M. L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging.* Springer-Verlag, New York.

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.

Ver Hoef, J. M., Cressie, N. and Barry, R. P. (2004). Flexible spatial models for kriging and cokriging using moving averages and the Fast Fourier Transform (FFT). *Journal of Computational and Graphical Statistics* **13**, 265–282.

Wackernagel, H. (2003). *Multivariate Geostatistics*, Berlin: Springer-Verlag, third edn.

Wang, B. and Xie, X. (1996). Low-frequency equatorial waves in vertically sheared zonal flow. Part I: stable waves. *Journal of the Atmospheric Sciences* **53**, 449–467.

Xie, X. and Wang, B. (1996). Low-frequency equatorial waves in vertically sheared zonal flow. Part II: unstable waves. *Journal of the Atmospheric Sciences* **53**, 3589–3605.

Zhang, H. and Zimmerman, D. L. (2005). Towards reconciling two asymptotic frameworks in spatial statistics. *Biometrika* **92**, 921–936.

Department of Applied Mathematics, University of Colorado, Boulder, CO 80309-0526, USA

E-mail: william.kleiber@colorado.edu