

# Modeling Tangential Vector Fields on a Sphere

## Abstract

Physical processes that manifest as tangential vector fields on a sphere are common in geophysical and environmental sciences. These naturally occurring vector fields are often subject to physical constraints, such as being curl-free or divergence-free. We start with constructing parametric models for curl-free and divergence-free vector fields that are tangential to the unit sphere through applying the surface gradient or the surface curl operator to a scalar random potential field on the unit sphere. **Using the Helmholtz-Hodge decomposition, we then construct a class of simple but flexible parametric models for general tangential vector fields, which are represented as a sum of a curl-free and a divergence-free components.** We propose a likelihood-based parameter estimation procedure, and show that fast computation is possible even for large data sets when the observations are on a regular latitude-longitude grid. Characteristics and practical utility of the proposed methodology are illustrated through extensive simulation studies and an application to a data set of ocean surface wind velocities collected by satellite-based scatterometers. We also compare our model with a bivariate Matérn model **and a non-stationary bivariate global model.**

**Keywords:** curl-free, divergence-free, Helmholtz-Hodge decomposition, Matérn model, ocean surface wind

# 1 Introduction

Vector fields defined on a spherical domain are principal objects of study in many branches of science. Terrestrial physical processes such as wind and oceanic currents, gravity, electric and magnetic fields are some of the most well-studied examples. In meteorology, the directionality of wind flow at surface level is an example of a tangential vector field on a sphere (Earth’s surface). Because the effective portion of the atmosphere and oceans are considerably thinner in their vertical extent in comparison with their horizontal extent, for many geophysical processes, the horizontal scale far exceeds the vertical scale. It is therefore natural to decompose such vector fields into tangential and radial components and treat them separately. For example, velocity divergence is expressed as a sum of horizontal (tangential) and vertical (radial) divergence in the continuity equation. Other such examples of vector fields include electric and magnetic fields in geophysics ([Sabaka et al., 2010](#)). A branch of solid-Earth geophysics focuses on the behavior of the Earth’s magnetic field, while its interactions with the solar wind are the focus of space-weather science. Electric and magnetic fields associated with ionospheric electric currents are distinct in the tangential direction from the radial direction, and many ionospheric electrodynamic processes can be treated as tangential vector fields on a sphere ([Richmond and Kamide, 1988](#)). In this paper, we focus on modeling the tangential component of a vector field on the unit sphere, especially its small-scale stochastic variations, which can form a significant part of the overall variability. Accounting for this variability can enhance model prediction accuracy, enable subgrid-scale inference, and facilitate better uncertainty quantification of model parameters. The modeling framework that we propose is flexible, incorporates small-scale variations, and is naturally interpretable and computationally tractable.

Gaussian random fields (GRF) have provided a very successful modeling framework for describing variations in many physical processes. Two key ingredients of the success of GRF modeling for stochastic processes are: (i) the behavior of the process is entirely characterized by the mean and the covariance functions, thereby facilitating deep theoretical investigations;

and (ii) computations for both estimation and prediction primarily involve matrix algebra. We provide a very brief overview of existing literature on the construction of covariance functions and random fields on a sphere. [Marinucci and Peccati \(2011\)](#) gave a detailed account of weakly isotropic spherical random fields, while [Guinness and Fuentes \(2016\)](#) broadly reviewed and compared isotropic covariance functions on spheres. [Hitczenko and Stein \(2012\)](#), [Jun and Stein \(2007\)](#), [Jun and Stein \(2008\)](#) mostly focused on the construction and characterization of **non-stationary** covariance functions for GRFs on a sphere.

Extensions of the modeling framework from scalar random fields to random vector fields pose an additional challenge due to the requirement of non-negative definiteness on the cross-covariance function of the vector field. Popular approaches to modeling vector fields include linear coregionalization models ([Bourgault and Marcotte, 1991](#); [Gelfand et al., 2004](#); [Goulard and Voltz, 1992](#)), multivariate Matérn models ([Apanasovich et al., 2012](#); [Gneiting et al., 2010](#)), and models derived from scalar potentials through differential operators ([Jun, 2011, 2014](#)); the latter being especially notable for their constructive approach. However, these models do not impose specific physical constraints on the Cartesian components of the vector field. There are many examples of vector fields satisfying “natural” physical constraints. For example, magnetic fields are solenoidal with zero divergence, and electric fields resulting from electric charges are curl-free. Vorticity, defined as the curl of a velocity vector field, plays an important role in atmospheric and oceanic dynamics in terms of characterizing the nature and degree of turbulence. Assuming water is incompressible, [Zhang et al. \(2007\)](#) represented lake water current velocity fields through applying the curl operator to a scalar stream function. [Constantinescu and Anitescu \(2013\)](#) developed cross-covariance models for vector fields that incorporate known physical constraints relating the behavior of their Cartesian components, and demonstrated that physics-based models can significantly outperform independence models in an application to geostrophic winds. These examples indicate that for observations of geophysical and environmental processes, kriging or data assimilation will benefit from modeling the corresponding vector fields as Gaussian processes

that incorporate physical constraints, such as being curl-free or divergence-free, into their construction.

We illustrate basic characteristics of a vector field with physical constraints by Figure 1, which displays an example of a pair of simulated divergence-free and curl-free random fields using our models that will be presented later in this paper. The details of simulating these vector fields are described in Section 3.1.

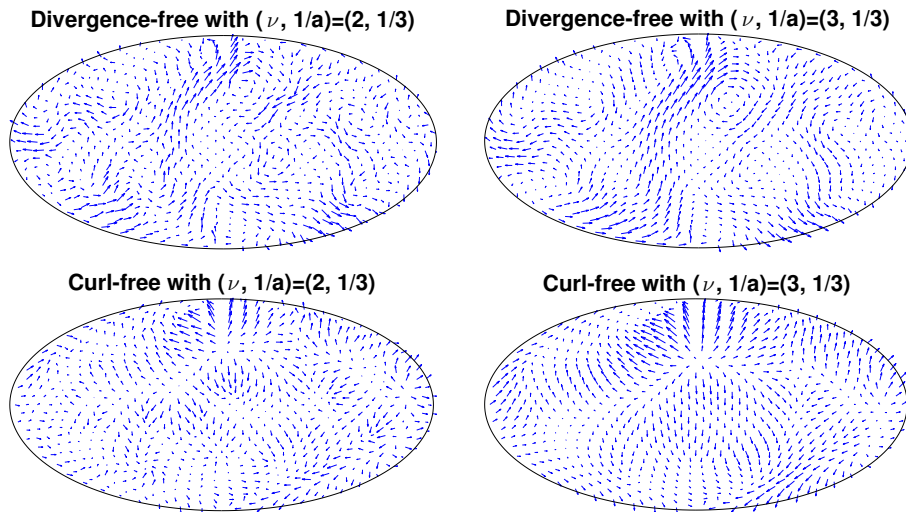


Figure 1: Simulated divergence-free and curl-free random fields on a HEALPix grid. The sphere has been projected to an ellipse by the Hammer projection. Using the same set of random deviates, these vector fields are simulated based on the Cholesky decomposition.

Modeling and analysis of vector fields (not necessarily tangential to a sphere) are facilitated by the celebrated *Helmholtz-Hodge decomposition* (Freedman and Schreiner, 2009; Kadri-Harouna and Perrier, 2012) that expresses the vector field as a sum of a curl-free component, a divergence-free component, and a component that is the gradient of a harmonic function. The harmonic component vanishes when the domain is a spherical surface and the vector field is tangential to the surface. In Section 2, we define the basic objects such as the surface gradient and the surface curl operators that are fundamental to the construction of a curl-free and a divergence-free components. Our work is related to that of Scheuerer and Schlather (2012), who constructed random vector fields that are either curl-free or divergence-free, and that of Schlather et al. (2015), who introduced a parametric

cross-covariance model for random vector fields consisting of the aforementioned curl-free and divergence-free components. However, these constructions are restricted to vector fields defined on Euclidean spaces only.

In this paper, we start with constructing Gaussian tangential vector fields on the unit sphere that are curl-free or divergence-free. Using the Helmholtz-Hodge decomposition, we then construct general Gaussian tangential vector fields from a pair of correlated Gaussian scalar potential fields. These constructions involve making careful use of spherical geometry and appropriate differential operators on the unit sphere. Through specifying a bivariate Matérn model for the pair of scalar potentials, we propose a simple but flexible parametric model named *Tangent Matérn Model (TMM)* with the following attractive characteristics:

- (i) The model does not introduce any additional parameters other than those already used in the bivariate Matérn model. The parameter of co-located correlation coefficient inherited from the bivariate Matérn model controls the correlation between the curl-free and the divergence-free components of the vector field, which is often non-zero for wind velocity fields (Cornford, 1998).
- (ii) The vector field can be equivalently represented in terms of its zonal and meridional components, and in this form, it allows for a negative correlation between two locations, which is common in meteorological variables (Daley, 1991, Chapter 4.3).
- (iii) The vector field is modeled as a sum of a curl-free and a divergence-free components. The magnitudes of the components, which are related to two important physical characteristics of the vector field, divergence and vorticity, can be inferred from parameter estimates.
- (iv) The cross-covariance structure of the vector field is not isotropic but axially symmetric.

We propose a likelihood-based estimation procedure for the TMM, and study its characteristics in terms of estimation through extensive numerical experiments. We also present

a fast computational algorithm that is applicable when the data are observed on a regular latitude-longitude grid. The proposed modeling framework, which is primarily used to describe the small-scale details of a vector field, is also extended to represent large-scale and spatio-temporal features.

As an illustration, we apply the proposed methodology to a data set of ocean surface wind velocities collected by the NASA’s Quick Scatterometer (QuikSCAT). Scatterometer data are important for numerical weather prediction (NWP), storm warning and monitoring (Atlas et al., 2001; Brennan et al., 2009). Over the past two decades, a number of statistical methods have been applied to model wind fields; for example, spatio-temporal Kalman filtering (Wikle and Cressie, 1999), spatio-temporal stationary covariance models (Cressie and Huang, 1999), spatio-temporal hierarchical Bayesian models (Wikle et al., 2001; Cressie and Wikle, 2011, Chapter 9.4), multivariate semiparametric Bayesian spatial models (Reich and Fuentes, 2007) and linear coregionalization models (Foley and Fuentes, 2008). In particular, using a planar approximation for the Earth’s surface, Cornford et al. (2004) fitted a version of the model of Schlather et al. (2015) derived from a Matérn covariance function to surface wind fields over a small sector of the north Atlantic ocean. Due to many-faceted complexity of the data, intricacies of available scientific models, and limitations of space, full-scale modeling of the wind velocity fields is beyond the scope of this paper. Instead, we focus here primarily on characterizing the residual wind fields after removing large-scale components through empirical orthogonal function (EOF) analysis, and demonstrating the implications of using the TMM to fit such fields. We also compare the predictive performance of co-kriging based on the TMM, a parsimonious bivariate Matérn model (Gneiting et al., 2010), and a non-stationary bivariate global model (Jun, 2011). The results reveal that the TMM has the best performance in estimation and competitive performance in prediction, which can be attributed to its ability of capturing certain physical characteristics of the wind fields.

The remainder of the paper is organized as follows. In Section 2, we present the construction of curl-free and divergence-free random tangential vector fields and the TMM. The

details of model fitting are also discussed in this section. We conduct extensive numerical experiments in Section 3 to explore the performance of the TMM. The proposed methodology is applied to a data set of QuikSCAT ocean surface wind velocities in Section 4. We discuss some relevant issues and future directions in Section 5. Details of theoretical derivations are collected in Supplementary Materials S.1.

## 2 Construction of Random Tangential Vector Fields

### 2.1 Surface Gradient and Surface Curl Operators

Tangential vector fields on the unit sphere can be constructed through applying appropriate differential operators to scalar potential fields. Let  $\mathbf{s}$  denote a point on the unit sphere  $\mathbb{S}^2 = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| = 1\}$  and  $(\theta, \phi)$  represent the same point in spherical coordinates, where  $\theta$  and  $\phi$  are the co-latitude and longitude, respectively. The tangent space at  $\mathbf{s} \equiv (\theta, \phi)$ , to be denoted by  $\mathcal{T}_{\mathbf{s}}$ , is a two-dimensional vector space with canonical orthonormal basis vectors  $\hat{\boldsymbol{\theta}}$  and  $\hat{\boldsymbol{\phi}}$ . These two vectors are defined such that  $\hat{\boldsymbol{\theta}}$  points southward, and is tangent to the great circle passing through the poles and  $\mathbf{s}$ , while  $\hat{\boldsymbol{\phi}}$  points eastward along the circle that is the intersection of  $\mathbb{S}^2$  with a plane, passing through  $\mathbf{s}$  and parallel to the equatorial plane. Besides, let  $\hat{\mathbf{r}}$  denote a unit vector at  $\mathbf{s}$  that points radially outward, i.e.,  $\hat{\mathbf{r}} = \hat{\boldsymbol{\theta}} \times \hat{\boldsymbol{\phi}}$ , where  $\times$  denotes the cross product on  $\mathbb{R}^3$ . All the vectors vary with  $\mathbf{s}$ , but for the moment we suppress this. In order to define spherical differential operators based on usual differential operators on  $\mathbb{R}^3$ , for any function or random field on  $\mathbb{S}^2$  mentioned hereafter, we always assume that it is actually defined on some spherical shell containing  $\mathbb{S}^2$ ,  $S_\epsilon = \{\mathbf{x} \in \mathbb{R}^3 : 1 - \epsilon < \|\mathbf{x}\| < 1 + \epsilon\}$ , where  $\epsilon > 0$ .

For the unit sphere, the surface gradient of a continuously differentiable function  $f : \mathbb{S}^2 \rightarrow \mathbb{R}$  is defined as the tangential component of its usual gradient, i.e.,

$$\nabla_{\mathbf{s}}^* f =: \mathbf{P}_{\mathbf{s}} \nabla_{\mathbf{s}} f,$$

where  $\mathbf{s} = (s_1, s_2, s_3)^T \in \mathbb{S}^2$ ,  $\mathbf{P}_{\mathbf{s}} = \mathbf{I}_3 - \mathbf{s}\mathbf{s}^T$  is the matrix that projects a vector in  $\mathbb{R}^3$  onto  $\mathcal{T}_{\mathbf{s}}$ , and  $\nabla_{\mathbf{s}}$  is the usual gradient. Then the surface curl of  $f$  is defined as a tangential vector field perpendicular to  $\nabla_{\mathbf{s}}^* f$ , i.e.,

$$L_{\mathbf{s}}^* f =: \hat{\mathbf{r}} \times \nabla_{\mathbf{s}}^* f = \mathbf{Q}_{\mathbf{s}} \nabla_{\mathbf{s}} f,$$

where the explicit expression of matrix  $\mathbf{Q}_{\mathbf{s}}$  is given in Supplementary Materials S.1.1.

If we further assume that  $f$  is twice continuously differentiable, then  $\nabla_{\mathbf{s}}^* f$  and  $L_{\mathbf{s}}^* f$  are curl-free and divergence-free, respectively, i.e.,  $\text{curl}_{\mathbf{s}}^*(\nabla_{\mathbf{s}}^* f) =: L_{\mathbf{s}}^* \cdot \nabla_{\mathbf{s}}^* f = 0$ , where  $\cdot$  denotes the dot (or scalar) product, and  $\text{div}_{\mathbf{s}}^*(L_{\mathbf{s}}^* f) =: \nabla_{\mathbf{s}}^* \cdot L_{\mathbf{s}}^* f = 0$ . If  $f$  is only continuously differentiable,  $\text{curl}_{\mathbf{s}}^*(\nabla_{\mathbf{s}}^* f)$  and  $\text{div}_{\mathbf{s}}^*(L_{\mathbf{s}}^* f)$  are not well-defined in the ordinary sense, and hence we call  $\nabla_{\mathbf{s}}^* f$  the *gradient-field* and  $L_{\mathbf{s}}^* f$  the *curl-field* derived from  $f$ . More details on differential calculus on the unit sphere can be found in [Freedden and Schreiner \(2009, Chapter 2\)](#). Some auxiliary notations and definitions are given in Supplementary Materials S.1.1.

## 2.2 Construction of Curl-free and Divergence-free Random Tangential Vector Fields

In this subsection, we present an approach for constructing curl-free and divergence-free random tangential vector fields on the unit sphere. They are derived through applying the surface gradient and the surface curl operators to a scalar random potential field, to be generically denoted by  $Z(\mathbf{s})$ . The cross-covariance structure of the resulting vector fields is determined by the covariance structure of the process  $Z$ . Suppose  $Z$  is defined on  $S_{\epsilon}$  with mean zero and finite variance, and satisfies the following regularity conditions:

- A1** The process  $Z(\mathbf{s})$  is differentiable in quadratic mean. Moreover, there exists a  $\mathbb{P}$ -a.e. sample continuously differentiable version of  $Z$ , denoted by  $\tilde{Z}$ , such that  $D^{(i)} \tilde{Z}(\mathbf{s}) = D_{\text{qm}}^{(i)} Z(\mathbf{s})$   $\mathbb{P}$ -a.e., where  $D^{(i)}$  and  $D_{\text{qm}}^{(i)}$  represent the sample partial derivative and the partial derivative in quadratic mean along the  $i$ -th coordinate direction, respectively.



**A2** The process  $Z(\mathbf{s})$  is stationary with twice continuously differentiable covariance function  $C(\mathbf{h})$ , where  $\mathbf{h}$  is the spatial separation vector between  $\mathbf{s}$  and  $\mathbf{t}$ , i.e.,  $\mathbf{h} = \mathbf{s} - \mathbf{t}$ .

The two regularity conditions ensure the validity of applying the differential operators. When  $Z(\mathbf{s})$  is Gaussian, **A1** can be verified through Theorems 3.2 and 4.3 in [Potthoff \(2010\)](#). By **A1**, there exists  $\Omega_0 \subset \Omega$  with  $\mathbb{P}(\Omega_0) = 1$  such that  $\tilde{Z}(\mathbf{s}, \omega)$  is continuously differentiable on  $S_\epsilon$  for any  $\omega \in \Omega_0$ . By applying the differential operators to  $\tilde{Z}$ , we construct two random tangential vector fields  $\mathbf{Y}_{\text{curl},Z}$  and  $\mathbf{Y}_{\text{div},Z}$  on  $\mathbb{S}^2$  such that  $\mathbf{Y}_{\text{curl},Z}(\mathbf{s}, \omega) = \mathbf{P}_\mathbf{s} \nabla_\mathbf{s} \tilde{Z}(\mathbf{s}, \omega)$  and  $\mathbf{Y}_{\text{div},Z}(\mathbf{s}, \omega) = \mathbf{Q}_\mathbf{s} \nabla_\mathbf{s} \tilde{Z}(\mathbf{s}, \omega)$  for any  $\omega \in \Omega_0$ . It is clear that the sample paths of  $\mathbf{Y}_{\text{curl},Z}$  and  $\mathbf{Y}_{\text{div},Z}$  are gradient-fields and curl-fields  $\mathbb{P}$ -a.e., respectively. If we further assume that the sample paths of  $\tilde{Z}$  are twice continuously differentiable  $\mathbb{P}$ -a.e., then  $\text{curl}_\mathbf{s}^*(\mathbf{Y}_{\text{curl},Z}(\mathbf{s}, \omega))$  and  $\text{div}_\mathbf{s}^*(\mathbf{Y}_{\text{div},Z}(\mathbf{s}, \omega))$  are well-defined and equal to zero for any  $\omega \in \Omega_0$ , and hence  $\mathbf{Y}_{\text{curl},Z}$  and  $\mathbf{Y}_{\text{div},Z}$  are curl-free and divergence-free, respectively.

The cross-covariance functions of  $\mathbf{Y}_{\text{curl},Z}$  and  $\mathbf{Y}_{\text{div},Z}$ , to be denoted by  $\mathbf{C}_{\text{curl},Z}$  and  $\mathbf{C}_{\text{div},Z}$ , respectively, are given explicitly in Theorem 1, the proof of which appears in Supplementary Materials S.1.2.

**Theorem 1.** *If **A1** and **A2** hold, then the cross-covariance functions  $\mathbf{C}_{\text{curl},Z}$  and  $\mathbf{C}_{\text{div},Z}$  can be represented as*

$$\mathbf{C}_{\text{curl},Z}(\mathbf{s}, \mathbf{t}) = -\mathbf{P}_\mathbf{s} \nabla_\mathbf{h} \nabla_\mathbf{h}^\text{T} C(\mathbf{h}) \Big|_{\mathbf{h}=\mathbf{s}-\mathbf{t}} \mathbf{P}_\mathbf{t}^\text{T}, \quad (1)$$

and

$$\mathbf{C}_{\text{div},Z}(\mathbf{s}, \mathbf{t}) = -\mathbf{Q}_\mathbf{s} \nabla_\mathbf{h} \nabla_\mathbf{h}^\text{T} C(\mathbf{h}) \Big|_{\mathbf{h}=\mathbf{s}-\mathbf{t}} \mathbf{Q}_\mathbf{t}^\text{T}. \quad (2)$$

We consider a special case of the construction described above, where  $Z(\mathbf{s})$  is an *isotropic* scalar random field on  $S_\epsilon$ . Then its covariance function  $C(\mathbf{h})$  can be written as  $C_1(\|\mathbf{h}\|)$  with  $C_1(r)$  being a function from  $[0, 2 + 2\epsilon)$  to  $\mathbb{R}$ . One popular choice for  $C_1$  is the Matérn model

$$M(r; \nu, a) = \frac{2^{1-\nu}}{\Gamma(\nu)} (ar)^\nu K_\nu(ar), \quad (3)$$

where  $K_\nu$  is the modified Bessel function of the second kind, the parameter  $\nu > 0$  controls the smoothness, and  $a > 0$  is the spatial scale parameter, whereby  $1/a$  controls the range of correlation. When  $Z(\mathbf{s})$  is assumed to be Gaussian with covariance function  $C_1(\|\mathbf{h}\|) = M(\|\mathbf{h}\|; \nu, a)$ ,  $\nu > 1$  is necessary and sufficient to ensure that conditions **A1** and **A2** hold. We give a proof of the statement in Supplementary Materials S.1.3, and an explicit expression of  $\nabla_{\mathbf{h}} \nabla_{\mathbf{h}}^T C_1(\|\mathbf{h}\|)$  in Appendix A.

We now establish connections of our proposal with existing literature. [Narcowich et al. \(2007\)](#) and [Fuselier and Wright \(2009\)](#) constructed a general class of positive definite curl-free and divergence-free kernels on  $\mathbb{S}^2$  using radial basis functions (RBFs) on  $\mathbb{R}^3$ . The cross-covariance kernels defined in (1) and (2) actually belong to this class. However, the constructions of [Narcowich et al. \(2007\)](#) and [Fuselier and Wright \(2009\)](#) are deterministic in nature, and the descriptions of their kernels do not automatically lead to the construction of random tangential vector fields. In contrast, our construction takes a different path, by directly deriving vector fields through the application of differential operators to scalar random potential fields. Moreover, we can extend the basic framework presented here to construct a rich class of spatio-temporal tangential vector fields on the unit sphere with more complicated cross-covariance structures and including large-scale components. These ideas will be developed in two stages, first by constructing a tangential vector field with correlated curl-free and divergence-free components from a pair of correlated scalar random potential fields (Section 2.3), and then by extending the framework to spatio-temporal processes (Section 2.5).

## 2.3 Tangent Matérn Model

In this subsection, we construct a class of parametric models for general Gaussian tangential vector fields using the Helmholtz-Hodge decomposition and the ingredients developed in Section 2.2. Let  $\mathbf{Z}(\mathbf{s}) = (Z_1(\mathbf{s}), Z_2(\mathbf{s}))^T$  denote an isotropic, bivariate, Gaussian random field on  $S_\epsilon$  with mean zero and cross-covariance function  $\mathbf{C}(\|\mathbf{h}\|) = \text{Cov}(\mathbf{Z}(\mathbf{s}), \mathbf{Z}(\mathbf{t}))$ , where

$\mathbf{h} = \mathbf{s} - \mathbf{t}$  and  $\mathbf{s}, \mathbf{t} \in S_\epsilon$ . The Helmholtz-Hodge decomposition states that any continuously differentiable tangential vector field on  $\mathbb{S}^2$  can be uniquely decomposed as a sum of a curl-free and a divergence-free components (Freedden and Schreiner, 2009). Following this idea, we first apply the spherical differential operators to  $Z_1$  and  $Z_2$  for obtaining a curl-free and a divergence-free tangential vector fields, respectively, as described in Section 2.2, and then define a tangential vector field as the sum of these two components. Specifically,

$$\mathbf{Y}_{\text{tan},\mathbf{Z}}(\mathbf{s}) = \mathbf{P}_s \nabla_s \tilde{Z}_1(\mathbf{s}) + \mathbf{Q}_s \nabla_s \tilde{Z}_2(\mathbf{s}) \quad \mathbb{P}\text{-a.e.}, \quad (4)$$

where  $\tilde{Z}_1$  and  $\tilde{Z}_2$  are  $\mathbb{P}$ -a.e. sample continuously differentiable versions of  $Z_1$  and  $Z_2$ , respectively. Since the transformation from  $\mathbf{Z}$  to  $\mathbf{Y}_{\text{tan},\mathbf{Z}}$  is linear, it is deduced that  $\mathbf{Y}_{\text{tan},\mathbf{Z}}$  is also Gaussian. This construction does not introduce any additional parameters, and allows for a correlation between the curl-free and the divergence-free components, which is inherited from the underlying bivariate potential field.

Gneiting et al. (2010) proposed a new class of cross-covariance functions called *multivariate Matérn* with a flexible correlation structure among its Cartesian components. We consider a *parsimonious bivariate Matérn model* for the underlying potential field  $\mathbf{Z}$  due to its simplicity and flexibility. “Parsimonious” here refers to the assumption that both Cartesian components of the bivariate process have the same spatial scale parameter. Gneiting et al. (2010) has argued that this assumption is not necessarily restrictive based on the fact that the parameters  $\sigma^2$  and  $a$  of a Matérn covariance function (with a fixed smoothness parameter  $\nu$ ) in dimension  $d \leq 3$  cannot be estimated consistently under infill asymptotics (Zhang, 2004). We name the resulting tangential vector field *Tangent Matérn Model (TMM)*. Write the cross-covariance function  $\mathbf{C}(\|\mathbf{h}\|)$  as  $(C_{ij}(\|\mathbf{h}\|))_{1 \leq i,j \leq 2}$ . The parsimonious bivariate Matérn model specifies

$$C_{ii}(\|\mathbf{h}\|) = \sigma_i^2 M(\|\mathbf{h}\|; \nu_i, a) \quad \text{for } i = 1, 2,$$

and

$$C_{12}(\|\mathbf{h}\|) = C_{21}(\|\mathbf{h}\|) = \rho_{12}\sigma_1\sigma_2 M(\|\mathbf{h}\|; (\nu_1 + \nu_2)/2, a),$$

where  $\sigma_i^2$  are the variance parameters, and  $a$  is the spatial scale parameter shared by the Cartesian components  $Z_1$  and  $Z_2$ . The parameter  $\rho_{12}$ , which represents a co-located correlation coefficient, controls the correlation between  $Z_1$  and  $Z_2$ , and through this, also determines the correlation between the curl-free and the divergence-free components of the tangential vector field  $\mathbf{Y}_{\text{tan},\mathbf{Z}}$ . A necessary and sufficient condition for non-negative definiteness of the cross-covariance function  $\mathbf{C}(\|\mathbf{h}\|)$  is that

$$|\rho_{12}| \leq \frac{\Gamma(\nu_1 + \frac{3}{2})^{1/2}}{\Gamma(\nu_1)^{1/2}} \frac{\Gamma(\nu_2 + \frac{3}{2})^{1/2}}{\Gamma(\nu_2)^{1/2}} \frac{\Gamma(\frac{1}{2}(\nu_1 + \nu_2))}{\Gamma(\frac{1}{2}(\nu_1 + \nu_2) + \frac{3}{2})}.$$

To ensure a sufficient degree of smoothness that enables the application of the differential operators, the smoothness parameters  $\nu_1$  and  $\nu_2$  are required to be larger than 1. Under these conditions, it can be shown that the derived cross-covariance function of  $\mathbf{Y}_{\text{tan},\mathbf{Z}}(\mathbf{s})$  is

$$\begin{aligned} & \mathbf{C}_{\text{tan},\mathbf{Z}}(\mathbf{s}, \mathbf{t}) \\ &= - \begin{pmatrix} \sigma_1 \mathbf{P}_s & \sigma_2 \mathbf{Q}_s \end{pmatrix} \begin{pmatrix} \mathbf{K}(\mathbf{h}; \nu_1, a) & \rho_{12} \mathbf{K}(\mathbf{h}; (\nu_1 + \nu_2)/2, a) \\ \rho_{12} \mathbf{K}(\mathbf{h}; (\nu_1 + \nu_2)/2, a) & \mathbf{K}(\mathbf{h}; \nu_2, a) \end{pmatrix} \begin{pmatrix} \sigma_1 \mathbf{P}_t^T \\ \sigma_2 \mathbf{Q}_t^T \end{pmatrix}, \quad (5) \end{aligned}$$

where  $\mathbf{K}(\mathbf{h}; \nu, a)$  denotes the RHS of (17) with  $F$  and  $G$  given by (20) and (21), respectively, and (17), (20) and (21) appear in Appendix A. Moreover, the cross-correlation between the curl-free and the divergence-free components

$$\text{Corr} \left( \mathbf{P}_s \nabla_s \tilde{Z}_1(\mathbf{s}), \mathbf{Q}_t \nabla_t \tilde{Z}_2(\mathbf{t}) \right) = -\rho_{12} \mathbf{\Lambda}(\mathbf{P}_s)^{-1/2} \mathbf{P}_s \mathbf{K}(\mathbf{h}; (\nu_1 + \nu_2)/2, a) \mathbf{Q}_t^T \mathbf{\Lambda}(\mathbf{Q}_t)^{-1/2},$$

where  $\mathbf{\Lambda}(\mathbf{P}_s) = \text{diag} \left( \text{diag} \left( -\mathbf{P}_s \mathbf{K}(\mathbf{0}; \nu_1, a) \mathbf{P}_s^T \right) \right)$ , and  $\mathbf{\Lambda}(\mathbf{Q}_t) = \text{diag} \left( \text{diag} \left( -\mathbf{Q}_t \mathbf{K}(\mathbf{0}; \nu_2, a) \mathbf{Q}_t^T \right) \right)$ .

We can relax the assumption on the underlying potential field  $\mathbf{Z}$  through specifying a full bivariate Matérn model (Gneiting et al., 2010) with distinct spatial scale parameters

and a more flexible cross-covariance smoothness parameter (see Supplementary Materials S.1.4). A more complicated constraint of the parameters is required to ensure the validity of the model, and hence the corresponding computation can be more expensive. Both the methodology and applications in the rest of this paper are focused on the TMM, but they can be extended to tangential vector fields derived from more general potential fields.

The tangential vector field  $\mathbf{Y}_{\text{tan},\mathbf{z}}(\mathbf{s})$  has been represented as a three-dimensional vector field in Cartesian coordinates so far. It can also be converted to a two-dimensional vector field  $\mathbf{V}(\mathbf{s})$  represented in the canonical coordinates  $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$  of the tangent space  $\mathcal{T}_{\mathbf{s}}$ , i.e.,

$$\mathbf{V}(\mathbf{s}) \equiv (u(\mathbf{s}), v(\mathbf{s}))^T = \mathbf{T}_{\mathbf{s}} \mathbf{Y}_{\text{tan},\mathbf{z}}(\mathbf{s}),$$

where  $\hat{\mathbf{u}} = \hat{\boldsymbol{\phi}}$ ,  $\hat{\mathbf{v}} = -\hat{\boldsymbol{\theta}}$ ,  $\mathbf{T}_{\mathbf{s}}$  is the transformation matrix with an explicit expression given in Supplementary Materials S.1.5, and  $u$  and  $v$  are called the zonal (eastward) and meridional (northward) components, respectively. The following proposition gives the cross-covariance function of  $\mathbf{V}$ , and its proof is deferred to Supplementary Materials S.1.5.

**Proposition 1.** *Let  $\mathbf{Y}_{\text{tan},\mathbf{z}}(\mathbf{s})$  be the resulting tangential vector field in the TMM defined by (4), where  $\nu_1, \nu_2 > 1$ , and  $\mathbf{V}(\mathbf{s}) \equiv (u(\mathbf{s}), v(\mathbf{s}))^T$  be its representation in the canonical coordinates  $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ . Then the cross-covariance function of  $\mathbf{V}$  is*

$$\mathbf{C}_{\mathbf{V}}(\mathbf{s}, \mathbf{t}) = \mathbf{T}_{\mathbf{s}} \mathbf{C}_{\text{tan},\mathbf{z}}(\mathbf{s}, \mathbf{t}) \mathbf{T}_{\mathbf{t}}^T. \quad (6)$$

When  $\mathbf{s} = \mathbf{t}$ ,

$$\mathbf{C}_{\mathbf{V}}(\mathbf{s}, \mathbf{s}) = -[\sigma_1^2 F_{\text{Mat}}(0; \nu_1, a) + \sigma_2^2 F_{\text{Mat}}(0; \nu_2, a)] \mathbf{I}_2. \quad (7)$$

The co-located cross-correlation between  $u$  and  $v$  is exactly zero, which is due to the isotropy of the underlying potential field.

We call a scalar random field  $X(\mathbf{s}) \equiv X(\theta_{\mathbf{s}}, \phi_{\mathbf{s}})$  on  $\mathbb{S}^2$  *axially symmetric* if

$$\text{Cov}(X(\theta_{\mathbf{s}}, \phi_{\mathbf{s}}), X(\theta_{\mathbf{t}}, \phi_{\mathbf{t}})) = C(\theta_{\mathbf{s}}, \theta_{\mathbf{t}}, \phi_{\mathbf{s}} - \phi_{\mathbf{t}}),$$

for any  $\theta_{\mathbf{s}}, \theta_{\mathbf{t}}, \phi_{\mathbf{s}}, \phi_{\mathbf{t}}$  and some function  $C$  (Jones, 1963). The following proposition gives an alternative representation of  $\mathbf{V}(\mathbf{s})$  when  $\mathbf{s}$  is represented as  $(\theta, \phi)$  in spherical coordinates, and its proof is deferred to Supplementary Materials S.1.6.

**Proposition 2.** *The vector field  $\mathbf{V}(\mathbf{s}) \equiv (u(\mathbf{s}), v(\mathbf{s}))^T$  can be represented (except at the two poles) as*

$$u(\theta, \phi) = \frac{1}{\sin \theta} \frac{\partial Z_1}{\partial \phi} + \frac{\partial Z_2}{\partial \theta} \quad \mathbb{P}\text{-a.e.}, \quad (8)$$

and

$$v(\theta, \phi) = \frac{1}{\sin \theta} \frac{\partial Z_2}{\partial \phi} - \frac{\partial Z_1}{\partial \theta} \quad \mathbb{P}\text{-a.e.}, \quad (9)$$

for any  $\theta \in (0, \pi)$ . Here the partial derivatives are defined in the sense of quadratic mean. It implies that  $u$  and  $v$  are axially symmetric both marginally and jointly.

The above representation suggests a similarity between the TMM and the non-stationary covariance and cross-covariance models proposed in Jun (2011, 2014); Jun and Stein (2008). All of them are represented as linear combinations of partial derivatives of scalar potential fields, and share the property of axial symmetry. However, they have important distinctions: (i) Jun's models are derived from a different perspective, which is to capture global non-stationarity with respect to latitude; and (ii) in Jun's models, the coefficients of the partial derivatives with respect to  $\theta$  and  $\phi$  are linear combinations of Legendre polynomials, i.e.,  $\sum_{j=0}^m a_j P_j(\sin(\pi/2 - \theta))$ , where  $P_j$  denotes the Legendre polynomial of order  $j$ , which introduce additional parameters  $a_j$  into the model. Jun (2011) assumes that  $Z_1$  and  $Z_2$  are independent, while Jun (2014) extends the model to allow for correlated  $Z_1$  and  $Z_2$  with spatially varying variance and smoothness parameters using the formulation of Kleiber and Nychka (2012).

## 2.4 Fast Parameter Estimation Using DFT

Suppose that  $\mathbf{V}(\mathbf{s}) = (u(\mathbf{s}), v(\mathbf{s}))^T$  is a bivariate Gaussian random field with mean zero and cross-covariance function (6). The observations on the process  $\mathbf{V}(\mathbf{s})$  (possibly with observational errors) at  $n$  different locations are collected in vector  $\mathbf{V} = (\mathbf{V}(\mathbf{s}_1)^T, \dots, \mathbf{V}(\mathbf{s}_n)^T)^T$ , which has a multivariate normal distribution of dimension  $2n$ . Henceforth, we use the notation  $\mathbf{V}$  to denote a random field and a random vector of observations interchangeably. The negative log-likelihood function (ignoring a constant) is

$$l(\boldsymbol{\theta}) = \frac{1}{2} \log |\boldsymbol{\Sigma}(\boldsymbol{\theta})| + \frac{1}{2} \mathbf{V}^T \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \mathbf{V}, \quad (10)$$

where  $\boldsymbol{\Sigma}(\boldsymbol{\theta})$  is the cross-covariance matrix and  $\boldsymbol{\theta}$  is the parameter vector. The method of maximum likelihood is commonly used to estimate parameters in spatial statistical models. However, the computation of the MLEs can be difficult for large data sets since the evaluation of the likelihood requires  $\mathcal{O}(n^3)$  operations. Nonetheless, in our case, when the observations are on a regular latitude-longitude grid, the discrete Fourier transform (DFT) can be used to speed up the computation. Regularly spaced observations are common for remote sensing satellite data, numerical weather model outputs and meteorological reanalyses. Jun (2011) pointed out that as long as the regular grid covers the full longitude range and the cross-covariance function is axially symmetric, the cross-covariance matrix can be transformed using the DFT to a  $2 \times 2$  block matrix with each block being a block diagonal matrix. The implementation details are given in Supplementary Materials S.1.7. The DFT helps reduce the time complexity of evaluating  $l(\boldsymbol{\theta})$  to  $\mathcal{O}(n(\log n_{\text{lon}} + n_{\text{lat}}^2))$ , where  $n_{\text{lat}}$  and  $n_{\text{lon}}$  denote the number of latitude and longitude grid points, respectively, and  $n_{\text{lat}} n_{\text{lon}} = n$ . A numerical comparison between the methods with and without using the DFT is given in Section 3.2. Note that the DFT is not suitable for irregularly spaced data nor data that are incomplete on a regular grid. In these cases, one may opt to use approximation methods such as covariance tapering (Furrer et al., 2006; Kaufman et al., 2008; Bevilacqua et al., 2015).

## 2.5 Spatio-temporal Modeling

The construction of spatial models for random tangential vector fields based on the spherical differential operators can be extended to spatio-temporal models. The idea is to start with a pair of correlated scalar potential fields on the unit sphere that vary in time. Let  $(\Phi, \Psi)$  denote the two scalar stochastic processes on  $\mathbb{S}^2 \times \mathbb{Z}$  with the following decompositions

$$\Phi(\mathbf{s}, t) = \mu_\Phi(\mathbf{s}) + \Phi_L(\mathbf{s}, t) + \Phi_S(\mathbf{s}, t); \quad \Psi(\mathbf{s}, t) = \mu_\Psi(\mathbf{s}) + \Psi_L(\mathbf{s}, t) + \Psi_S(\mathbf{s}, t), \quad (11)$$

where  $\mu_\Phi$  and  $\mu_\Psi$  are nonrandom spatial means,  $(\Phi_L, \Psi_L)$  and  $(\Phi_S, \Psi_S)$  are large-scale and small-scale spatio-temporal components, respectively. It is often assumed that large-scale components are independent of small-scale ones. The large-scale components are modeled by finite linear combinations of nonrandom spatially varying basis functions, with random coefficients that are functions of time, or through certain parametric stochastic processes of  $(\mathbf{s}, t)$ . The former is given by

$$\Phi_L(\mathbf{s}, t) = \sum_{j=1}^J a_j(t) \phi_j(\mathbf{s}); \quad \Psi_L(\mathbf{s}, t) = \sum_{k=1}^K b_k(t) \psi_k(\mathbf{s}), \quad (12)$$

where  $\{\phi_j\}_{j=1}^J$  and  $\{\psi_k\}_{k=1}^K$  are nonrandom basis functions, and  $\{a_j(t)\}_{j=1}^J$  and  $\{b_k(t)\}_{k=1}^K$  are zero-mean time series, typically assumed to be stationary. Such models have been used for modeling spatio-temporal processes defined on both Euclidean and spherical domains. Possible choices for the basis functions are spherical harmonics ([Stein, 2007](#)), wavelets ([Matsuo et al., 2011](#); [Nychka et al., 2002](#)) and overcomplete frames of functions ([Hsu et al., 2012](#); [Nychka et al., 2015](#)). For the small-scale components, we may assume that the bivariate process  $(\Phi_S(\mathbf{s}, t), \Psi_S(\mathbf{s}, t))$  is independent across time, and for any fixed  $t$ , it is an isotropic, bivariate, Gaussian random field with mean zero and the same cross-covariance structure.

For simplicity, we may assume that the scalar potential fields are detrended, i.e.,  $\mu_\Phi \equiv 0$



and  $\mu_\Psi \equiv 0$ . Applying the spherical differential operators to  $\Phi$  and  $\Psi$ , we have

$$\nabla_{\mathbf{s}}^* \Phi(\mathbf{s}, t) = \sum_{j=1}^J a_j(t) \tilde{\phi}_j(\mathbf{s}) + \nabla_{\mathbf{s}}^* \Phi_S(\mathbf{s}, t); \quad L_{\mathbf{s}}^* \Psi(\mathbf{s}, t) = \sum_{k=1}^K b_k(t) \tilde{\psi}_k(\mathbf{s}) + L_{\mathbf{s}}^* \Psi_S(\mathbf{s}, t), \quad (13)$$

where  $\tilde{\phi}_j(\mathbf{s}) = \nabla_{\mathbf{s}}^* \phi_j(\mathbf{s})$  and  $\tilde{\psi}_k(\mathbf{s}) = L_{\mathbf{s}}^* \psi_k(\mathbf{s})$ . Thus, the resulting vector field is

$$\mathbf{Y}(\mathbf{s}, t) = \sum_{j=1}^J a_j(t) \tilde{\phi}_j(\mathbf{s}) + \sum_{k=1}^K b_k(t) \tilde{\psi}_k(\mathbf{s}) + \nabla_{\mathbf{s}}^* \Phi_S(\mathbf{s}, t) + L_{\mathbf{s}}^* \Psi_S(\mathbf{s}, t).$$

Representing  $\mathbf{Y}(\mathbf{s}, t)$  in the canonical coordinates  $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$  of the tangent space  $\mathcal{T}_{\mathbf{s}}$ , we have

$$\mathbf{V}(\mathbf{s}, t) = \mathbf{T}_{\mathbf{s}} \mathbf{Y}(\mathbf{s}, t) = \sum_{j=1}^J a_j(t) \mathbf{T}_{\mathbf{s}} \tilde{\phi}_j(\mathbf{s}) + \sum_{k=1}^K b_k(t) \mathbf{T}_{\mathbf{s}} \tilde{\psi}_k(\mathbf{s}) + \mathbf{T}_{\mathbf{s}} (\nabla_{\mathbf{s}}^* \Phi_S(\mathbf{s}, t) + L_{\mathbf{s}}^* \Psi_S(\mathbf{s}, t)). \quad (14)$$

The sum of the first two terms forms the large-scale component of the vector field, and the last term represents its small-scale component. For any fixed  $t$ , the last term is a tangential vector field analogous to the TMM.

## 3 Numerical Experiments

### 3.1 Simulation of Random Fields

Figure 1 displays simulated zero-mean divergence-free (first row) and curl-free (second row) tangential vector fields from the models constructed in Section 2.2. The scalar potential field  $Z(\mathbf{s})$  follows the Matérn model  $M(\|\mathbf{h}\|; \nu, a)$ , and the sampling locations are on a HEALPix grid (Górski et al., 2005) with 768 grid points, which partitions the unit sphere into equal area pixels. The smoothness parameters are 2 and 3 for the realizations on the left and right, respectively. The spatial scale parameter  $a = 3$ . As we can see from Figure 1, the realizations on the left are rougher than those on the right, which conforms with the parameter specification.

### 3.2 Accuracy of Parameter Estimation

In this subsection, we conduct Monte Carlo simulations to evaluate the accuracy of parameter estimation for the TMM. Suppose that  $\mathbf{V}(\mathbf{s})$  is a zero-mean bivariate Gaussian random field that follows the TMM (6). Augmented with nugget effects to account for observational errors, the cross-covariance function hence becomes

$$\mathbf{C}_{\mathbf{V}}(\mathbf{s}, \mathbf{t}) = \mathbf{T}_{\mathbf{s}} \mathbf{C}_{\text{tan}, \mathbf{Z}}(\mathbf{s}, \mathbf{t}) \mathbf{T}_{\mathbf{t}}^{\text{T}} + \text{diag} \left( \tau_1^2 \mathbf{1}(\mathbf{s} = \mathbf{t}), \tau_2^2 \mathbf{1}(\mathbf{s} = \mathbf{t}) \right), \quad (15)$$

where the parameter vector  $\boldsymbol{\theta} = (\sigma_1, \sigma_2, \rho_{12}, \nu_1, \nu_2, 1/a, \tau_1, \tau_2)$ . To reduce the computational burden and mimic the TOMS Level 3 data analyzed in Jun and Stein (2008), the sampling locations are on a regular latitude-longitude grid with latitudes ranging from  $-50$  degree to  $50$  degree.

We implement the DFT described in Section 2.4 to evaluate the negative log-likelihood function (10), and compare its computational speed with that of the method without using the DFT. On a regular laptop with a 2.4 GHz Intel Core i5 processor, when the number of latitudes  $n_{\text{lat}} = 25$  and longitudes  $n_{\text{lon}} = 50$ , the method using the DFT only takes 3.03 seconds, while the one without using the DFT takes 34.43 seconds. Due to its capability of handling large-scale problems, we use the *interior-point* algorithm through the Matlab function *fmincon* to minimize the negative log-likelihood function. The details of initial value specification are given in Supplementary Materials S.1.8. Note that the range of the smoothness parameters is assumed to be  $(1, 5]$  since too large values would be unrealistic and result in numerical instability. From our extensive numerical experience, the algorithm typically converges after 100 to 200 iterations. A further computational gain can be achieved using parallel computing in gradient estimation at each iteration.

We conduct 500 simulation runs, and generate realizations using the cross-covariance function (15) with  $\boldsymbol{\theta} = (1, 1, 0.5, 3, 4, 1/2, 0.1, 0.1)$  on three regular grids with  $(n_{\text{lat}}, n_{\text{lon}}) = (10, 20), (15, 30)$  and  $(20, 40)$ , respectively. The sample size is increased to check the behavior

of the MLEs from the perspective of infill asymptotics (i.e., fixed domain asymptotics). We have also conducted some smaller-scale experiments on irregular grids, which yield similar results (see Supplementary Materials S.2.1). Figure 2 displays the boxplots of the MLEs. For all the parameters, the medians of the estimates are close to their true values, i.e., the estimates have small biases. The spreads of the estimates for  $\nu_2$  are considerably larger than those for  $\nu_1$ . This can be explained by the following two facts. First, it is more difficult to estimate the smoothness parameter with a larger value, which conforms with the findings in Gneiting et al. (2010) and references therein. Moreover, (7) implies that the ratio between the variability of the curl-free and the divergence-free components is  $\sigma_1^2(\nu_2 - 1)/(\sigma_2^2(\nu_1 - 1))$ , and hence larger than 1 given  $\sigma_1 = \sigma_2$  and  $\nu_1 < \nu_2$ . This suggests that the data contain more information on the curl-free component than the divergence-free one. As we have previously mentioned, the parameters  $\sigma^2$  and  $a$  of a Matérn covariance function (with a fixed smoothness parameter  $\nu$ ) in dimension  $d \leq 3$  cannot be consistently estimated under infill asymptotics. Since the TMM is constructed based on a bivariate Matérn model, it is not surprising that the spreads of the estimates for  $\sigma_1$  and  $\sigma_2$  do not decrease significantly when the sample size increases. Additional numerical results can be found in Supplementary Materials S.2.2-S.2.4.

The standard errors of the MLEs can be estimated using a parametric bootstrap procedure, whose effectiveness is investigated in Supplementary Materials S.2.5.

### 3.3 Predictive Performance Comparison

The predictive performance of the TMM is compared with the parsimonious bivariate Matérn model (PARS-BM) (see Section 2.3), and the non-stationary bivariate global model (NBG) proposed in Jun (2011). In the NBG, a zero-mean bivariate random field  $\mathbf{V}(\mathbf{s}) = (V_1(\mathbf{s}), V_2(\mathbf{s}))^T$  is specified such that

$$V_i(\theta, \phi) = \left\{ a_i \frac{\partial}{\partial \theta} + b_i \frac{\partial}{\partial \phi} \right\} Z(\theta, \phi) + U_i(\theta, \phi) \quad \text{for } i = 1, 2,$$

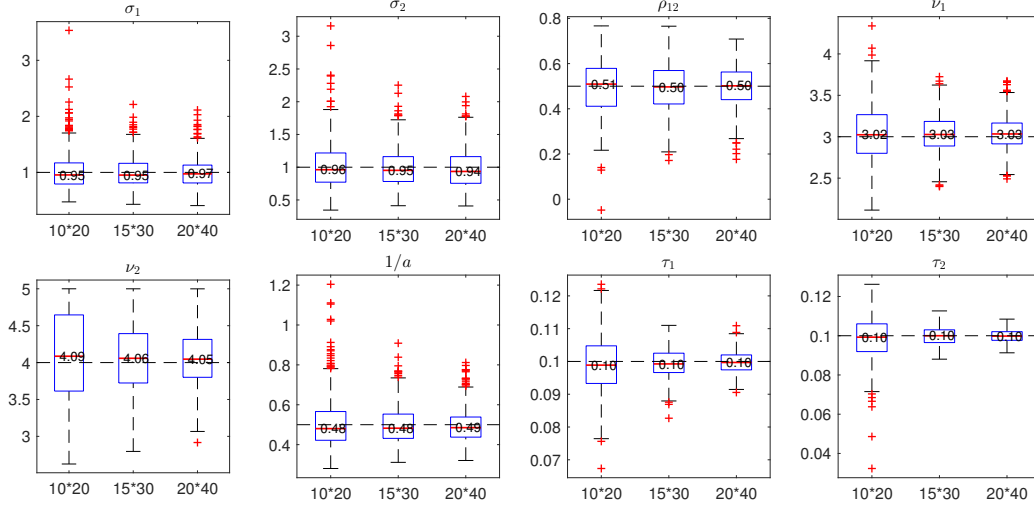


Figure 2: Results of the Monte Carlo simulation study to investigate the accuracy of parameter estimation for the TMM. The MLEs of  $\theta$  are summarized by boxplots for three simulated data sets with increasing sample size (shown on x-axis). The dashed horizontal lines are at the true values.

where  $Z$  follows the unscaled Matérn model  $2^{\kappa-1}\Gamma(\kappa)M(r;\kappa,a) = (ar)^{\kappa}K_{\kappa}(ar)$  with  $r = \|\mathbf{h}\|$ ,  $\mathbf{U} = (U_1, U_2)^T$  follows the PARS-BM with  $\theta = (\sigma_1, \sigma_2, \rho_{12} \equiv 0, \nu_1, \nu_2, 1/a, \tau_1, \tau_2)$ , and  $Z, U_1$  and  $U_2$  are independent. We further assume  $a_1 > 0$  to avoid identifiability problems. The results are reported in Supplemental Materials S.2.6, in which the TMM has the best predictive performance among the three models.

## 4 Data Example

In this section, we illustrate the effectiveness of the proposed TMM and the associated methodology through applying them to the data set of QuikSCAT ocean surface wind velocities introduced in Section 1. There are two versions of the data set: Level 2B and Level 3, which are available on <http://podaac.jpl.nasa.gov>, and a comprehensive description of them is given in the user's manual [Piolle and Bentamy \(2002\)](#). The Level 3 version is used here, which contains monthly mean ocean surface winds from January 2000 through December 2008. We focus on modeling the horizontal components of the ocean surface winds, and

hence the observations are the zonal and meridional winds in [m/s], abbreviated hereafter as  $u$  and  $v$  winds. The sampling locations are on an (incomplete) regular grid with spatial resolution 1 degree latitude by 1 degree longitude.

The horizontal component of the wind field is naturally a tangential vector field, which exhibits variability over both space and time. An extensive scientific literature (e.g., [Shukla and Saha, 1974](#); [Bijlsma et al., 1986](#)) suggests modeling the horizontal component  $\mathbf{Y}(\mathbf{s}, t)$ , represented in Cartesian coordinates, by a velocity potential  $\Phi$  and a stream function  $\Psi$ , i.e.,

$$\mathbf{Y}(\mathbf{s}, t) = \nabla_{\mathbf{s}}^* \Phi(\mathbf{s}, t) + L_{\mathbf{s}}^* \Psi(\mathbf{s}, t) = \mathbf{P}_{\mathbf{s}} \nabla_{\mathbf{s}} \Phi(\mathbf{s}, t) + \mathbf{Q}_{\mathbf{s}} \nabla_{\mathbf{s}} \Psi(\mathbf{s}, t),$$

where  $\mathbf{s} \in \mathbb{S}^2$ ,  $t$  is a time index, and  $\Phi$  and  $\Psi$  are **scalar processes**. The importance of the above representation as a sum of a curl-free and a divergence-free components can be gauged from the fact that the vorticity, entirely determined by the divergence-free component, is a key ingredient in the analysis and prediction of cyclonic events ([Holton, 2004](#), Chapter 4.2). Note that the velocity potential and stream function are allowed to be correlated, which has been justified in [Cornford \(1998\)](#). We impose certain structures on the velocity potential and stream function, and specifically use the spatio-temporal model described in Section 2.5.

The true wind field is typically unobservable. Suppose that the observations  $\mathbf{V}(\mathbf{s}_i, t_j) = (V_1(\mathbf{s}_i, t_j), V_2(\mathbf{s}_i, t_j))^T$ ,  $i = 1, \dots, N, j = 1, \dots, T$  satisfy

$$\mathbf{V}(\mathbf{s}_i, t_j) = \mathbf{T}_{\mathbf{s}} \mathbf{Y}(\mathbf{s}_i, t_j) + \boldsymbol{\epsilon}(\mathbf{s}_i, t_j), \quad (16)$$

where  $V_1$  and  $V_2$  correspond to the  $u$  and  $v$  winds, respectively,  $N$  is assumed to be larger than  $T$ , and  $\boldsymbol{\epsilon}(\mathbf{s}_i, t_j)$  are observational errors, which are modeled as i.i.d.  $\mathcal{N}(\mathbf{0}, \text{diag}(\tau_1^2, \tau_2^2))$ . Due to limitations of space, and in the interest of a focused analysis, our goal here is primarily to demonstrate the effectiveness of the proposed methodology in describing the small-scale component of the wind field, rather than to conduct a detailed statistical analysis of the full-scale wind field. We subtract a crude estimate of the large-scale component from the

wind field as a way of extracting the small-scale component. The large-scale component is estimated from the data by the method of vector empirical orthogonal functions (VEOF) (Pan et al., 2001, 2003). Let  $\mathbf{V} = (\mathbf{V}_1, \mathbf{V}_2)$  denote the matrix of observations, where  $\mathbf{V}_1$  and  $\mathbf{V}_2$  are  $T \times N$  matrices corresponding to the  $u$  and  $v$  winds, respectively. We center  $\mathbf{V}$  by subtracting the vector of column averages from each row. The singular value decomposition (SVD) is applied to  $\mathbf{V}$ , i.e.,

$$(\mathbf{V}_1, \mathbf{V}_2) = \mathbf{V} = \mathbf{U}\mathbf{D}\mathbf{W}^T = \mathbf{U}\mathbf{D}(\mathbf{W}_1^T, \mathbf{W}_2^T),$$

where  $\mathbf{U}$  is a  $T \times T$  orthogonal matrix with  $u_k(t_j)$  in row  $j$ , column  $k$ ,  $\mathbf{D} = \text{diag}(d_1, \dots, d_T)$  with singular values  $d_1 \geq d_2 \geq \dots \geq d_T \geq 0$ , and  $\mathbf{W}$  is a  $2N \times T$  matrix with orthonormal columns, i.e., VEOFs. Element-wisely, the observations can be expressed as  $\mathbf{V}(\mathbf{s}_i, t_j) = \sum_{k=1}^T d_k u_k(t_j) (w_{1,k}(\mathbf{s}_i), w_{2,k}(\mathbf{s}_i))^T$ , where  $w_{1,k}(\mathbf{s}_i)$  and  $w_{2,k}(\mathbf{s}_i)$  are the elements of  $\mathbf{W}_1$  and  $\mathbf{W}_2$  in row  $i$ , column  $k$ , respectively. The cumulative sum of the ordered squared singular values indicates that the first 64 VEOFs can explain approximately 95% of the overall variability in the observed wind fields. It is not surprising to obtain 64 VEOFs since the VEOF method is applied to the  $u$  and  $v$  winds simultaneously over the entire globe. We use this as a crude measure of the large-scale variability in the data, and hence approximate the large-scale component by  $\sum_{k=1}^{64} d_k u_k(t_j) (w_{1,k}(\mathbf{s}_i), w_{2,k}(\mathbf{s}_i))^T$ . The residuals after subtracting the estimated large-scale component are regarded as the small-scale component, corrupted by the observational errors.

To reduce the computational burden and for a quick comparison of various models, we choose a subregion of the Indian Ocean (57.30°E–114.59°E, 57.30°S–28.65°N) with spatial resolution 2 degree latitude by 2 degree longitude. The subregion is large enough (the range of latitudes is almost 90°) so that models treating the domain as a subset of the Euclidean plane cannot handle the distortion caused by the curvature. Using a coarser grid may lead to a loss of information, but the parameter estimates are comparable to those for the original grid

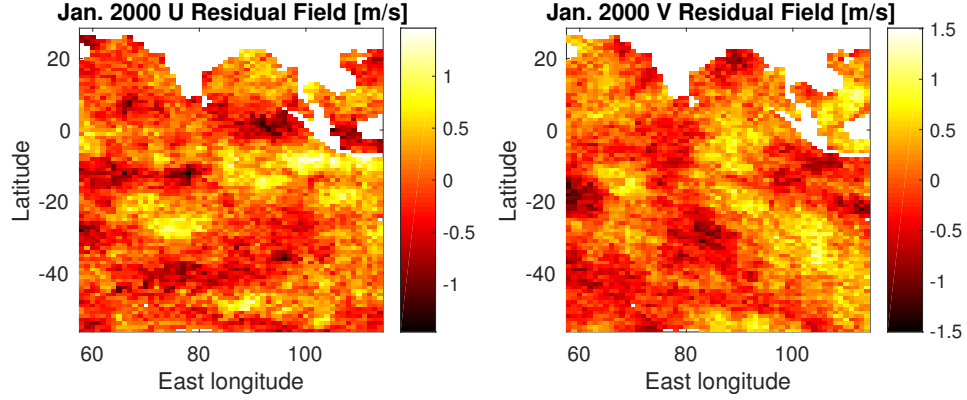


Figure 3: An example of the  $u$  and  $v$  residual wind fields for January, 2000 in the subregion of the Indian Ocean after subtracting the first 64 VEOFs.

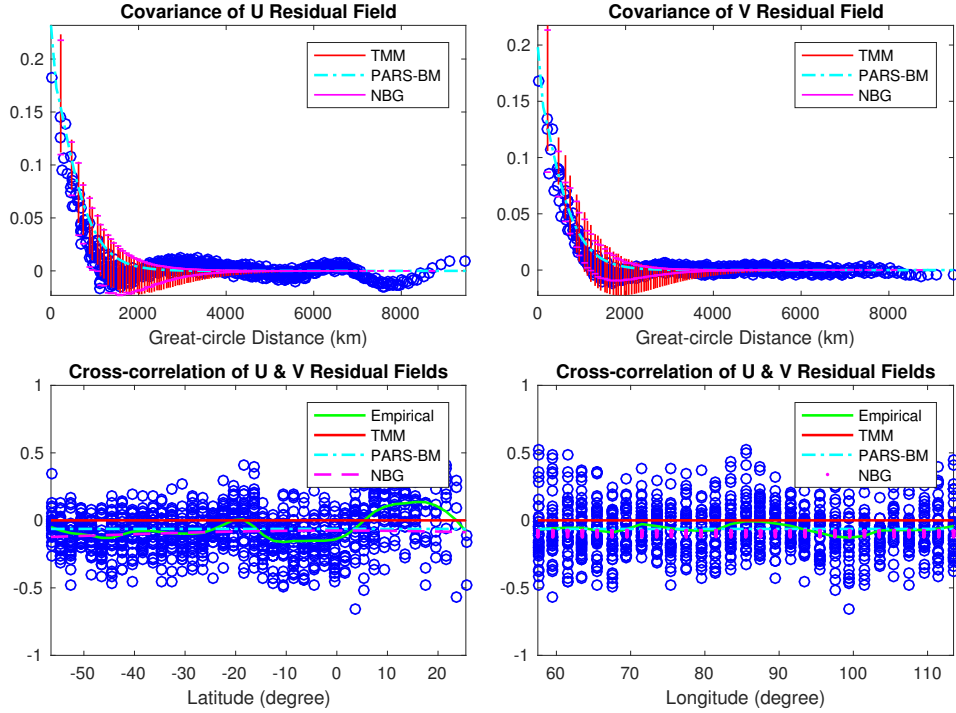


Figure 4: **First row:** Empirical and fitted covariances of the  $u$  and  $v$  residual fields. Bin medians of the empirical covariances are shown by circles. The spreads (min to max) of the fitted covariances by the TMM and NBG (due to anisotropy) at each great-circle distance are summarized by a vertical line and two horizontal lines at the min and max values, respectively. Dash-dot lines are for the PARS-BM. **Second row:** Empirical and fitted co-located cross-correlations between the  $u$  and  $v$  residual fields, i.e.,  $\text{Corr}(u(s), v(s))$ : circles for the empirical co-located cross-correlations, green solid lines for the loess curves fitted to the empirical co-located cross-correlations, red solid lines for the TMM, dash-dot lines for the PARS-BM, and dashed lines for the NBG. They are plotted with respect to latitude (left) and longitude (right).

according to our numerical experience (e.g., the estimated smoothness parameters become slightly larger). Note that there are 1070 observations for each month, and hence  $108 \times 1070$  observations in total, which are assumed to be independent across months. The sample autocorrelation functions, and the marginal and the Chi-Square Q-Q plots suggest that the assumptions of temporal independence and Gaussianity on the  $u$  and  $v$  residual wind fields are reasonable (see Supplementary Materials S.3.1). Figure 3 contains an example of the  $u$  and  $v$  residual wind fields for January, 2000 in the subregion of the Indian Ocean.

Table 1: Maximum likelihood estimates of the parameters for the TMM, PARS-BM and NBG, applied to the  $u$  and  $v$  residual wind fields from January 2000 through December 2008 in the subregion of the Indian Ocean. Their bootstrap standard errors are shown in parentheses. “-” means the parameter is not available for the corresponding model

Model	TMM	PARS-BM	NBG
$\sigma_1$	0.029 (4.25e-4)	0.429 (3.14e-3)	0.270
$\sigma_2$	0.055 (8.05e-4)	0.396 (3.09e-3)	0.343
$\rho_{12}$	0.281 (7.05e-3)	-0.080 (6.88e-3)	-
$\nu_1$	1.758 (0.022)	1.239 (0.038)	0.721
$\nu_2$	2.034 (0.020)	1.132 (0.035)	0.830
$1/a$	0.106 (1.80e-3)	0.058 (1.45e-3)	0.069
$a_1$	-	-	0.016
$a_2$	-	-	-0.0046
$b_1$	-	-	0.0026
$b_2$	-	-	0.012
$\kappa$	-	-	2.867
$\tau_1$	0.210 (1.48e-3)	0.218 (1.37e-3)	0.213
$\tau_2$	0.196 (1.48e-3)	0.203 (1.61e-3)	0.195
Log-likelihood	-45126	-46995	-45137
# of parameters	8	8	12

We first fit the TMM to the  $u$  and  $v$  residual wind fields for all the months by minimizing the negative joint log-likelihood function. Without loss of generality, the Earth is regarded as a unit sphere. The MLEs and their bootstrap standard errors are displayed in Table 1, where the latter are computed using the parametric bootstrap procedure (see Supplementary Materials S.2.5) with 200 bootstrap samples. Recall that the ratio between the variability of the divergence-free and the curl-free components is  $\sigma_2^2(\nu_1 - 1)/(\sigma_1^2(\nu_2 - 1)) = 2.637$ , which



suggests that the former is the dominant component of the wind field. This result can be explained by considering a purely divergence-free wind field, called geostrophic wind. It is the theoretical wind that results from an exact force balance between the Coriolis effect and the pressure gradient force. Surface winds that blow over the oceans are close to being geostrophic because of the relatively smooth ocean surface (Park, 2001, page 283). Moreover, the estimated co-located correlation coefficient is significantly different from zero, which justifies the necessity of allowing the curl-free and the divergence-free components to be correlated. Note that the signal-to-noise ratios of the  $u$  and  $v$  residual wind fields are  $\text{Var}(u(\mathbf{s}))/\tau_1^2 \equiv 4.040$  and  $\text{Var}(v(\mathbf{s}))/\tau_2^2 \equiv 4.653$ , respectively.

For comparison, we also fit the PARS-BM using the R package RandomFields (Version 3.0.62), and the NBG (see Section 3.3), where the computational details of the latter can be found in Supplementary Materials S.1.9. The MLEs and their bootstrap standard errors are displayed in Table 1. The TMM achieves the highest likelihood with the fewest number of parameters, while the NBG has a slightly lower likelihood but adds four additional parameters. With the same number of parameters as the TMM, the PARS-BM yields a considerably lower likelihood than the other two models. The fits of the TMM and NBG suggest rough  $u$  and  $v$  residual wind fields with effective smoothness around 0.7 to 0.8, while that of the PARS-BM favors smoother fields with  $\nu_1 = 1.239$  and  $\nu_2 = 1.132$ .

Figure 4 shows the empirical and fitted covariances and co-located cross-correlations, i.e.,  $\text{Corr}(u(\mathbf{s}), v(\mathbf{s}))$ , of the  $u$  and  $v$  residual wind fields under the TMM, PARS-BM and NBG. The trends of the fitted covariances are generally in agreement with the empirical ones. Compared with the PARS-BM, which is restricted to non-negative covariances, the TMM and NBG capture the negative empirical covariances around the great-circle distance of 2000 km, which are a well-accepted characteristic in meteorological variables (Daley, 1991, Chapter 4.3). According to Proposition 1, the fitted co-located cross-correlations under the TMM are identically zero. The estimated co-located correlation coefficient under the PARS-BM ( $\rho_{12} = -0.080$ ) leads to fitted co-located cross-correlations with a constant

small negative value. They are close to those under the NBG, although the latter exhibit mild non-stationarity with respect to latitude. Due to the subtraction of the estimated large-scale component, the empirical co-located cross-correlations have trends moderately oscillating around zero, which seem to conform with the fits. The trends would significantly deviate from zero if fewer VEOFs (i.e., less large-scale variability) were subtracted.

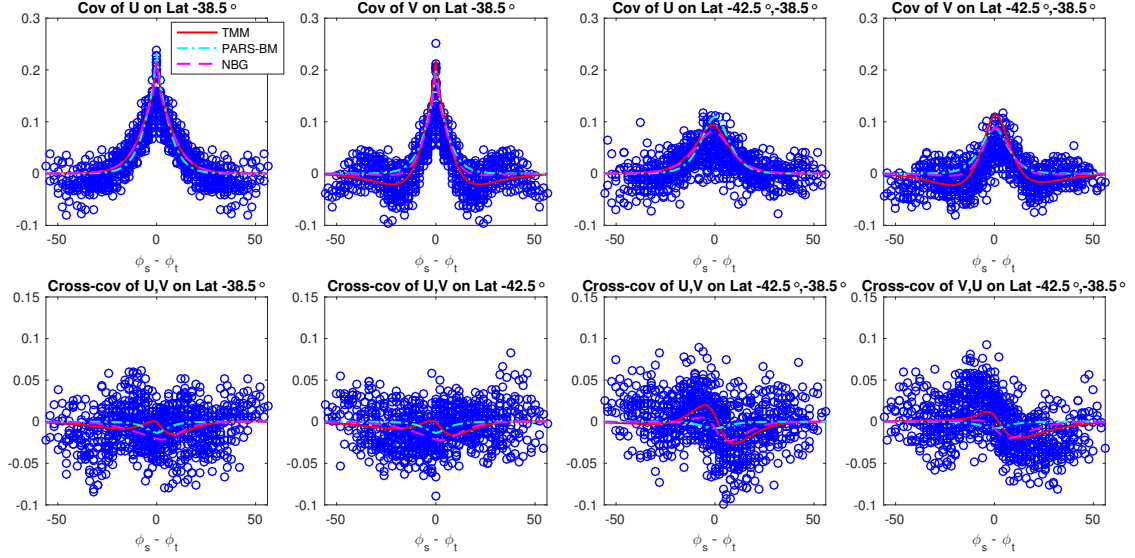


Figure 5: **First row:** Empirical and fitted covariances of the  $u$  and  $v$  residual fields, represented as a function of  $\theta_s$ ,  $\theta_t$  and  $\phi_s - \phi_t$ . **Second row:** Empirical and fitted cross-covariances between the  $u$  and  $v$  residual fields, represented as a function of  $\theta_s$ ,  $\theta_t$  and  $\phi_s - \phi_t$ . Circles are for the empirical covariances or cross-covariances, solid lines are for the TMM, dash-dot lines are for the PARS-BM, and dashed lines are for the NBG.

To examine the three models from the perspective of axial symmetry, Figure 5 displays the empirical and fitted covariances and cross-covariances of the  $u$  and  $v$  residual wind fields, as a function of  $\theta_s$ ,  $\theta_t$  and  $\phi_s - \phi_t$  for certain latitudes. The fits capture the peaks of the empirical covariances around  $\phi_s = \phi_t$ , but only those of the TMM match well with the negative values on the tails of the empirical covariances of the  $v$  residual fields. Some of the trends of the empirical cross-covariances, which are asymmetric with respect to  $\phi_s = \phi_t$ , are depicted by the fits of the TMM and NBG, but missed by those of the PARS-BM. A comparison between the empirical and fitted variances of the  $u$  and  $v$  residual wind fields is given in Supplementary Materials S.3.2.

Table 2: Cokriging cross-validation scores on the  $u$  and  $v$  residual wind fields, averaged over 20 cross-validation replications for the TMM, PARS-BM and NBG. The standard deviations of the scores over the 20 replications are shown in parentheses

Model	Variable	MSPE	MAE	LogS	CRPS
TMM	$u$	0.1126 (0.0020)	0.2587 (0.0021)	0.2933 (0.0078)	0.1845 (0.0015)
	$v$	0.1034 (0.0017)	0.2495 (0.0020)	0.2484 (0.0084)	0.1772 (0.0014)
PARS-BM	$u$	0.1191 (0.0020)	0.2650 (0.0022)	0.3153 (0.0079)	0.1891 (0.0015)
	$v$	0.1062 (0.0015)	0.2531 (0.0016)	0.2595 (0.0074)	0.1795 (0.0012)
NBG	$u$	0.1109 (0.0019)	0.2572 (0.0021)	0.2878 (0.0078)	0.1833 (0.0015)
	$v$	0.1025 (0.0016)	0.2491 (0.0018)	0.2475 (0.0078)	0.1768 (0.0012)

Finally, we compare the predictive performance of the TMM, PARS-BM and NBG. To estimate the parameters, we randomly select half of the 1070 locations outside a  $20^\circ$  (width) by  $40^\circ$  (height) rectangular region in the center of the subregion of the Indian Ocean, which are consistent over all the time points. Then at each time point, we perform cokriging to predict at the remaining locations using the parameter estimates just obtained. Four popular scoring rules: the mean squared prediction error (MSPE), the mean absolute error (MAE), the logarithmic score (LogS), and the continuous ranked probability score (CRPS) (Gneiting and Raftery, 2007) are used to assess the prediction accuracy, averaged over all the predicted locations and time points. We repeat the cross-validation procedure 20 times, and compute the means and standard deviations (in parentheses) of the resulting scores, which are displayed in Table 2. The NBG yields the best predictive performance, which is slightly better than the TMM in terms of the values of the scores. This is not surprising because both models include a differential operator term, which has been shown to play an important role in capturing certain characteristics of the data, and the former is a more complicated model with four additional parameters. It has also been found in Jun (2011) that lower prediction errors are associated with more complicated cross-covariance models. To assess the chance variability in the results, we conduct a series of two-sample  $t$ -tests with the Bonferroni correction. They suggest that the differences in the scores of the TMM and

NBG are not statistically significant, but both of them significantly outperform the PARS-BM (see Supplementary Materials S.3.3). A comparison among the wind fields predicted by the three models and the observations is given in Supplementary Materials S.3.4.

## 5 Discussion

A limitation of the Tangent Matérn Model is the assumption of isotropy for the underlying potential field. This assumption leads to Proposition 1, in which the co-located cross-covariance function is constant over the sphere, and in particular,  $\text{Corr}(u(\mathbf{s}), v(\mathbf{s})) = 0$ . Nonetheless, we can extend the model using an anisotropic underlying potential field, such as non-stationary bivariate Matérn models with spatially varying parameters (Kleiber and Nychka, 2012; Jun, 2014). The spatially varying parameters can be modeled by parametric functions of spherical harmonics (Bolin and Lindgren, 2011) or covariates observed together with the vector field (Risser and Calder, 2015). This remains a topic for future research.

Proposition 2 reveals the connection of our model with those of Jun (2011, 2014); Jun and Stein (2008), although they are derived from different perspectives. Also, as demonstrated by Proposition 2, the tangential vector field that follows our model is axially symmetric when represented in spherical coordinates. This finding enables fast computation for large data sets when the observations are on a regular latitude-longitude grid. For example, on average, it takes approximately 8 hours (on a machine with a 2.60GHz Intel Xeon E5-2690 v3 processor) to estimate the parameters when the sample size is as large as 5000. For irregularly spaced observations, apart from the method of covariance tapering mentioned in Section 2.4, a multi-resolution model constructed by the Wendland radial basis functions can also be used to represent or approximate the underlying potential field with a Matérn or a more complicated covariance structure (Nychka et al., 2015). Since the key matrices involved in the computation are sparse, this model can be applied to large data sets.

## Appendix A Isotropic $Z(\mathbf{s})$

Since  $C_1(\|\mathbf{h}\|)$  is twice continuously differentiable when  $h_1 \geq 0$  and  $h_2 = h_3 = 0$ ,  $C_1(r)$  is also twice continuously differentiable. Moreover, the fact that  $C_1(\|\mathbf{h}\|)$  is continuously differentiable at  $\mathbf{h} = \mathbf{0}$  implies that  $C_1'(0) = 0$ . Then we have

$$\nabla_{\mathbf{h}} \nabla_{\mathbf{h}}^T C_1(\|\mathbf{h}\|) = \begin{cases} F(\|\mathbf{h}\|) \mathbf{I}_3 + G(\|\mathbf{h}\|) \mathbf{h} \mathbf{h}^T & \text{if } \mathbf{h} \neq \mathbf{0} \\ F(0) \mathbf{I}_3 & \text{if } \mathbf{h} = \mathbf{0}, \end{cases} \quad (17)$$

where

$$F(r) = \begin{cases} \frac{1}{r} C_1'(r) & \text{if } r > 0 \\ C_1''(0) & \text{if } r = 0, \end{cases} \quad (18)$$

and

$$G(r) = \frac{1}{r} \left( \frac{1}{r} C_1'(r) \right)' \quad \text{if } r > 0. \quad (19)$$

Here the value of  $\nabla_{\mathbf{h}} \nabla_{\mathbf{h}}^T C_1(\|\mathbf{h}\|)$  at  $\mathbf{h} = \mathbf{0}$  is obtained by

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} F(\|\mathbf{h}\|) \mathbf{I}_3 + G(\|\mathbf{h}\|) \mathbf{h} \mathbf{h}^T = F(0) \mathbf{I}_3.$$

When  $C_1$  is chosen as the Matérn model with  $\nu > 1$ , we have explicit expressions of  $F$  and  $G$ , i.e.,

$$F_{\text{Mat}}(r; \nu, a) = \begin{cases} -\frac{2^{1-\nu}}{\Gamma(\nu)} a^2 (ar)^{\nu-1} K_{\nu-1}(ar) & \text{if } r > 0 \\ -\frac{a^2}{2(\nu-1)} & \text{if } r = 0, \end{cases} \quad (20)$$

and

$$G_{\text{Mat}}(r; \nu, a) = \frac{2^{1-\nu}}{\Gamma(\nu)} a^4 (ar)^{\nu-2} K_{\nu-2}(ar) \quad \text{if } r > 0. \quad (21)$$

## References

- 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(497):180–193.
- Atlas, R., Hoffman, R. N., Leidner, S. M., Sienkiewicz, J., Yu, T.-W., Bloom, S. C., Brin, E., Ardizzone, J., Terry, J., Bungato, D., and Jusem, J. C. (2001). The effects of marine winds from scatterometer data on weather analysis and forecasting. *Bulletin of the American Meteorological Society*, 82:1965–1990.
- Bevilacqua, M., Fassò, A., Gaetan, C., Porcu, E., and Velandia, D. (2015). Covariance tapering for multivariate Gaussian random fields estimation. *Statistical Methods & Applications*, pages 1–17.
- Bijlsma, S., Hafkenscheid, L., and Lynch, P. (1986). Computation of the streamfunction and velocity potential and reconstruction of the wind field. *Monthly weather review*, 114(8):1547–1551.
- Bolin, D. and Lindgren, F. (2011). Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. *The Annals of Applied Statistics*, 5(1):523–550.
- Bourgault, G. and Marcotte, D. (1991). Multivariable variogram and its application to the linear model of coregionalization. *Mathematical Geology*, 23(7):899–928.
- Brennan, M. J., Hennon, C. C., and Knabb, R. D. (2009). The operational use of QuikSCAT ocean surface vector winds at the National Hurricane Center. *Weather and Forecasting*, 24:621–645.
- Constantinescu, E. M. and Anitescu, M. (2013). Physics-based covariance models for Gaussian processes with multiple outputs. *International Journal for Uncertainty Quantification*, 3(1).
- Cornford, D. (1998). Flexible Gaussian process wind field models. Technical Report NCRG/98/017, Neural Computing Research Group, Aston University, Birmingham.
- Cornford, D., Csató, L., Evans, D. J., and Oppen, M. (2004). Bayesian analysis of the scatterometer wind retrieval inverse problem: some new approaches. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 66(3):609–626.
- Cressie, N. and Huang, H.-C. (1999). Classes of nonseparable, spatio-temporal stationary covariance functions. *Journal of the American Statistical Association*, 94(448):1330–1339.
- Cressie, N. and Wikle, C. K. (2011). *Statistics for spatio-temporal data*. John Wiley & Sons.
- Daley, R. (1991). *Atmospheric data analysis*. Cambridge: Cambridge University Press.

- Foley, K. and Fuentes, M. (2008). A statistical framework to combine multivariate spatial data and physical models for hurricane surface wind prediction. *Journal of Agricultural, Biological, and Environmental Statistics*, 13(1):37–59.
- Freeden, W. and Schreiner, M. (2009). *Spherical functions of mathematical geosciences: a scalar, vectorial, and tensorial setup*. Advances in Geophysical and Environmental Mechanics and Mathematics. Springer-Verlag Berlin Heidelberg.
- Furrer, R., Genton, M. G., and Nychka, D. (2006). Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics*, 15(3):502–523.
- Fuselier, E. J. and Wright, G. B. (2009). Stability and error estimates for vector field interpolation and decomposition on the sphere with RBFs. *SIAM Journal on Numerical Analysis*, 47(5):3213–3239.
- Gelfand, A. E., Schmidt, A. M., Banerjee, S., and Sirmans, C. F. (2004). Nonstationary multivariate process modeling through spatially varying coregionalization. *Test*, 13(2):263–312.
- Gneiting, T., Kleiber, W., and Schlather, M. (2010). Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 105(491):1167–1177.
- Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378.
- Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., and Bartelmann, M. (2005). HEALPix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere. *The Astrophysical Journal*, 622(2):759.
- Goulard, M. and Voltz, M. (1992). Linear coregionalization model: Tools for estimation and choice of cross-variogram matrix. *Mathematical Geology*, 24(3):269–286.
- Guinness, J. and Fuentes, M. (2016). Isotropic covariance functions on spheres: Some properties and modeling considerations. *Journal of Multivariate Analysis*, 143:143–152.
- Hitczenko, M. and Stein, M. L. (2012). Some theory for anisotropic processes on the sphere. *Statistical Methodology*, 9:211–227.
- Holton, J. R. (2004). *An introduction to dynamic meteorology*. International geophysics series. Elsevier Academic Press, Amsterdam, Boston, Heidelberg.
- Hsu, N.-J., Chang, Y.-M., and Huang, H.-C. (2012). A group lasso approach for non-stationary spatial-temporal covariance estimation. *Environmetrics*, 23(1):12–23.
- Jones, R. H. (1963). Stochastic processes on a sphere. *Annals of Mathematical Statistics*, 34(1):213–218.
- Jun, M. (2011). Non-stationary cross-covariance models for multivariate processes on a globe. *Scandinavian Journal of Statistics*, 38(4):726–747.

- Jun, M. (2014). Matérn-based nonstationary cross-covariance models for global processes. *Journal of Multivariate Analysis*, 128:134–146.
- Jun, M. and Stein, M. L. (2007). An approach to producing space-time covariance functions on spheres. *Technometrics*, 49(4):468–479.
- Jun, M. and Stein, M. L. (2008). Nonstationary covariance models for global data. *The Annals of Applied Statistics*, 2(4):1271–1289.
- Kadri-Harouna, S. and Perrier, V. (2012). Helmholtz-Hodge decomposition on  $[0, 1]^d$  by divergence-free and curl-free wavelets. In Boissonnat, J.-D., Chenin, P., Cohen, A., Gout, C., Lyche, T., Mazure, M.-L., and Schumaker, L., editors, *Curves and Surfaces*, volume 6920 of *Lecture Notes in Computer Science*, pages 311–329. Springer Berlin Heidelberg.
- Kaufman, C. G., Schervish, M. J., and Nychka, D. W. (2008). Covariance tapering for likelihood-based estimation in large spatial data sets. *Journal of the American Statistical Association*, 103(484):1545–1555.
- Kleiber, W. and Nychka, D. (2012). Nonstationary modeling for multivariate spatial processes. *Journal of Multivariate Analysis*, 112:76–91.
- Marinucci, D. and Peccati, G. (2011). *Random Fields on the Sphere: Representation, Limit Theorems and Cosmological Applications*, volume 389 of *London Mathematical Society Lecture Note Series*. Cambridge University Press.
- Matsuo, T., Nychka, D. W., and Paul, D. (2011). Nonstationary covariance modeling for incomplete data: Monte carlo em approach. *Computational Statistics & Data Analysis*, 55(6):2059–2073.
- Narcowich, F. J., Ward, J. D., and Wright, G. B. (2007). Divergence-free RBFs on surfaces. *Journal of Fourier Analysis and Applications*, 13(6):643–663.
- Nychka, D., Bandyopadhyay, S., Hammerling, D., Lindgren, F., and Sain, S. (2015). A multiresolution gaussian process model for the analysis of large spatial datasets. *Journal of Computational and Graphical Statistics*, 24(2):579–599.
- Nychka, D., Wikle, C., and Royle, J. A. (2002). Multiresolution models for nonstationary spatial covariance functions. *Statistical Modelling*, 2(4):315–331.
- Pan, J., Yan, X.-H., Zheng, Q., and Liu, W. T. (2001). Vector empirical orthogonal function modes of the ocean surface wind variability derived from satellite scatterometer data. *Geophysical Research Letters*, 28(20):3951–3954.
- Pan, J., Yan, X.-H., Zheng, Q., Liu, W. T., and Klemas, V. V. (2003). Interpretation of scatterometer ocean surface wind vector EOFs over the northwestern Pacific. *Remote sensing of environment*, 84(1):53–68.
- Park, C. C. (2001). *The environment: principles and applications*. Psychology Press.



- Piolle, J.-F. and Bentamy, A. (2002). QuikSCAT scatterometer mean wind field products user manual. *Ifremer, Department of Oceanography from Space, Ref.: C2-MUT-W-04-IF, Version 1.0, May 2002.*
- Potthoff, J. (2010). Sample properties of random fields III: Differentiability. *Communications on Stochastic Analysis*, 4(3):335–353.
- Reich, B. J. and Fuentes, M. (2007). A multivariate semiparametric bayesian spatial modeling framework for hurricane surface wind fields. *The Annals of Applied Statistics*, 1(1):249–264.
- Richmond, A. D. and Kamide, Y. (1988). Mapping electrodynamic features of the high-latitude ionosphere from localized observations: Technique. *Journal of Geophysical Research: Space Physics*, 93(A6):5741–5759.
- Risser, M. D. and Calder, C. A. (2015). Regression-based covariance functions for nonstationary spatial modeling. *Environmetrics*, 26(4):284–297.
- Sabaka, T. J., Hulot, G., and Olsen, N. (2010). Mathematical properties relevant to geomagnetic field modeling. In Freedon, W., Nashed, M. Z., and Sonar, T., editors, *Handbook of Geomathematics*, pages 503–538. Springer Berlin Heidelberg.
- Scheuerer, M. and Schlather, M. (2012). Covariance models for divergence-free and curl-free random vector fields. *Stochastic Models*, 28(3):433–451.
- Schlather, M., Malinowski, A., Menck, P. J., Oesting, M., and Strokorb, K. (2015). Analysis, simulation and prediction of multivariate random fields with package RandomFields. *Journal of Statistical Software*, 63(8):1–25.
- Shukla, J. and Saha, K. (1974). Computation of non-divergent streamfunction and irrotational velocity potential from the observed winds. *Monthly weather review*, 102(6):419–425.
- Stein, M. L. (2007). Spatial variation of total column ozone on a global scale. *The Annals of Applied Statistics*, pages 191–210.
- Wikle, C. K. and Cressie, N. (1999). A dimension-reduced approach to space-time Kalman filtering. *Biometrika*, 86(4):815–829.
- Wikle, C. K., Milliff, R. F., Nychka, D., and Berliner, L. M. (2001). Spatiotemporal hierarchical bayesian modeling tropical ocean surface winds. *Journal of the American Statistical Association*, 96(454):382–397.
- Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *Journal of the American Statistical Association*, 99(465):250–261.
- Zhang, Z., Beletsky, D., Schwab, D. J., and Stein, M. L. (2007). Assimilation of current measurements into a circulation model of Lake Michigan. *Water resources research*, 43(11).