

Supplementary Materials for “Modeling Tangential Vector Fields on a Sphere”

S.1 Additional Technical Details

S.1.1 Auxiliary Notations and Definitions

In this subsection, we describe some auxiliary notations and definitions helpful for understanding the construction and characteristics of tangential vector fields on the unit sphere.

The matrix \mathbf{Q}_s is defined as

$$\mathbf{Q}_s =: \hat{\mathbf{r}} \times \mathbf{P}_s = \begin{pmatrix} 0 & -s_3 & s_2 \\ s_3 & 0 & -s_1 \\ -s_2 & s_1 & 0 \end{pmatrix}. \quad (\text{S.1})$$

The surface gradient and the surface curl operators can be represented in spherical coordinates except at the two poles, i.e.,

$$\nabla_s^* f = \frac{\partial f}{\partial \theta} \hat{\boldsymbol{\theta}} + \frac{1}{\sin \theta} \frac{\partial f}{\partial \phi} \hat{\boldsymbol{\phi}},$$

and

$$L_s^* f = \frac{\partial f}{\partial \theta} \hat{\boldsymbol{\phi}} - \frac{1}{\sin \theta} \frac{\partial f}{\partial \phi} \hat{\boldsymbol{\theta}}.$$

Let $X(\mathbf{s})$ denote a scalar random field on a spherical shell S_ϵ , where $(\Omega, \mathcal{F}, \mathbb{P})$ is the underlying probability space. The partial derivative of $X(\mathbf{s})$ in quadratic mean along the i -th coordinate

direction is defined as a stochastic process $D_{\text{qm}}^{(i)}X(\mathbf{s})$ such that

$$\mathbb{E} \left(\frac{X(\mathbf{s} + h\mathbf{e}_i) - X(\mathbf{s})}{h} - D_{\text{qm}}^{(i)}X(\mathbf{s}) \right)^2 \rightarrow 0 \quad \text{as } h \rightarrow 0,$$

where \mathbf{e}_i is the unit vector in \mathbb{R}^3 with the i -th element one and zeros elsewhere. The sample partial derivatives of $X(\mathbf{s})$, $D^{(i)}X(\mathbf{s})$, $i = 1, 2, 3$, are defined as the usual partial derivatives of $X(\mathbf{s}, \omega)$ for any fixed $\omega \in \Omega$. A random field \tilde{X} is said to be a \mathbb{P} -a.e. sample continuously differentiable version of X if the following hold:

- (i) For any $\mathbf{s} \in S_\epsilon$, $\mathbb{P} \left(\tilde{X}(\mathbf{s}) = X(\mathbf{s}) \right) = 1$.
- (ii) There exists $\Omega_0 \subset \Omega$ with $\mathbb{P}(\Omega_0) = 1$ such that $\tilde{X}(\mathbf{s}, \omega)$ is continuously differentiable on S_ϵ for any $\omega \in \Omega_0$.

A random tangential vector field on \mathbb{S}^2 is said to be curl-free (or divergence-free) if its sample paths are curl-free (or divergence-free) \mathbb{P} -a.e.

S.1.2 Proof of Theorem 1

We first verify (1). Since $D^{(i)}\tilde{Z}(\mathbf{s}) = D_{\text{qm}}^{(i)}Z(\mathbf{s})$ \mathbb{P} -a.e., $\mathbf{Y}_{\text{curl},Z}(\mathbf{s}) = \mathbf{P}_\mathbf{s} \nabla_{\mathbf{s},\text{qm}} Z(\mathbf{s})$ \mathbb{P} -a.e., where $\nabla_{\mathbf{s},\text{qm}}$ is the usual gradient on \mathbb{R}^3 defined in the sense of quadratic mean. Thus,

$$\begin{aligned} \mathbf{C}_{\text{curl},Z}(\mathbf{s}, \mathbf{t}) &= \mathbf{P}_\mathbf{s} \text{Cov}(\nabla_{\mathbf{s},\text{qm}} Z(\mathbf{s}), \nabla_{\mathbf{t},\text{qm}} Z(\mathbf{t})) \mathbf{P}_\mathbf{t}^\text{T} \\ &= \mathbf{P}_\mathbf{s} [\text{Cov}(D_{\text{qm}}^{(i)}Z(\mathbf{s}), D_{\text{qm}}^{(j)}Z(\mathbf{t}))]_{1 \leq i, j \leq 3} \mathbf{P}_\mathbf{t}^\text{T} \\ &= \mathbf{P}_\mathbf{s} \left[\frac{\partial^2 C}{\partial s_i \partial t_j}(\mathbf{s} - \mathbf{t}) \right]_{1 \leq i, j \leq 3} \mathbf{P}_\mathbf{t}^\text{T} \quad (\text{by condition A2}) \\ &= -\mathbf{P}_\mathbf{s} \left[\frac{\partial^2 C}{\partial h_i \partial h_j}(\mathbf{h}) \right]_{1 \leq i, j \leq 3} \bigg|_{\mathbf{h}=\mathbf{s}-\mathbf{t}} \mathbf{P}_\mathbf{t}^\text{T} \\ &= -\mathbf{P}_\mathbf{s} \nabla_\mathbf{h} \nabla_\mathbf{h}^\text{T} C(\mathbf{h}) \bigg|_{\mathbf{h}=\mathbf{s}-\mathbf{t}} \mathbf{P}_\mathbf{t}^\text{T}. \end{aligned}$$

(2) can be derived similarly.

S.1.3 Proof That $\nu > 1$ Is Necessary and Sufficient

First, $\nu > 1$ is necessary because the underlying potential field $Z(\mathbf{s})$ is required to be differentiable in quadratic mean. Second, we shall show the sufficiency. Since $\nu > 1$ and Z is Gaussian, the sample paths of Z are continuously differentiable \mathbb{P} -a.e. (Handcock and Wallis, 1994). Besides, Z has partial derivatives in quadratic mean. These two facts together imply that $(Z(\mathbf{s} + h\mathbf{e}_i) - Z(\mathbf{s})) / h$ converges to $D_{\text{qm}}^{(i)}Z(\mathbf{s})$ in L_2 (and in probability), and converges to $D^{(i)}Z(\mathbf{s})$ almost everywhere (and in probability). By the uniqueness of convergence in probability, $D_{\text{qm}}^{(i)}Z(\mathbf{s}) = D^{(i)}Z(\mathbf{s})$ \mathbb{P} -a.e. Thus, we can choose Z as \tilde{Z} so that condition **A1** is satisfied. Jun and Stein (2007) pointed out that the Matérn model (3) is twice continuously differentiable and $M'(0; \nu, a) = 0$. This implies that $C_1(\|\mathbf{h}\|)$ is also twice continuously differentiable when $\|\mathbf{h}\| > 0$. It can be shown by definition that

$$\frac{\partial C_1(\|\mathbf{h}\|)}{\partial h_i} = \begin{cases} C_1'(\|\mathbf{h}\|) \frac{h_i}{\|\mathbf{h}\|} & \text{if } \mathbf{h} \neq \mathbf{0} \\ 0 & \text{if } \mathbf{h} = \mathbf{0}. \end{cases}$$

Based on the first derivatives, we can also show by definition that

$$\left. \frac{\partial^2 C_1(\|\mathbf{h}\|)}{\partial h_i \partial h_j} \right|_{\mathbf{h}=\mathbf{0}} = \begin{cases} 0 & \text{if } i \neq j \\ C_1'''(0) & \text{if } i = j. \end{cases}$$

This is equivalent to

$$\left. \nabla_{\mathbf{h}} \nabla_{\mathbf{h}}^T C_1(\|\mathbf{h}\|) \right|_{\mathbf{h}=\mathbf{0}} = F(0) \mathbf{I}_3.$$

Besides, we know that

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \left. \nabla_{\mathbf{h}} \nabla_{\mathbf{h}}^T C_1(\|\mathbf{h}\|) \right|_{\mathbf{h} \neq \mathbf{0}} = \lim_{\mathbf{h} \rightarrow \mathbf{0}} F(\|\mathbf{h}\|) \mathbf{I}_3 + G(\|\mathbf{h}\|) \mathbf{h} \mathbf{h}^T = F(0) \mathbf{I}_3.$$

Thus, $C_1(\|\mathbf{h}\|)$ is twice continuously differentiable at $\mathbf{h} = \mathbf{0}$, i.e., condition **A2** is satisfied.

S.1.4 Tangential Vector Fields Derived from a Full Bivariate Matérn Model

The underlying potential field $\mathbf{Z}(\mathbf{s})$ is assumed to follow a full bivariate Matérn model (Gneiting et al., 2010), which specifies

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

and

$$C_{12}(\|\mathbf{h}\|) = C_{21}(\|\mathbf{h}\|) = \rho_{12} \sigma_1 \sigma_2 M(\|\mathbf{h}\|; \nu_{12}, a_{12}).$$

It is valid if and only if

$$\rho_{12}^2 \leq \frac{\Gamma(\nu_1 + \frac{d}{2})}{\Gamma(\nu_1)} \frac{\Gamma(\nu_2 + \frac{d}{2})}{\Gamma(\nu_2)} \frac{\Gamma(\nu_{12})^2}{\Gamma(\nu_{12} + \frac{d}{2})^2} \frac{a_1^{2\nu_1} a_2^{2\nu_2}}{a_{12}^{4\nu_{12}}} \inf_{t \geq 0} \frac{(a_{12}^2 + t^2)^{2\nu_{12} + d}}{(a_1^2 + t^2)^{\nu_1 + (d/2)} (a_2^2 + t^2)^{\nu_2 + (d/2)}}. \quad (\text{S.2})$$

Condition (S.2) implies that if $\nu_{12} < \frac{1}{2}(\nu_1 + \nu_2)$, ρ_{12} has to be zero, i.e., the two Cartesian components of \mathbf{Z} are independent. Thus, our primary interest lies in the case where $\nu_{12} \geq \frac{1}{2}(\nu_1 + \nu_2)$.

The derived cross-covariance function of $\mathbf{Y}_{\text{tan}, \mathbf{Z}}(\mathbf{s})$ is

$$\begin{aligned} & \mathbf{C}_{\text{tan}, \mathbf{Z}}(\mathbf{s}, \mathbf{t}) \\ &= - \begin{pmatrix} \sigma_1 \mathbf{P}_{\mathbf{s}} & \sigma_2 \mathbf{Q}_{\mathbf{s}} \end{pmatrix} \begin{pmatrix} \mathbf{K}(\mathbf{h}; \nu_1, a_1) & \rho_{12} \mathbf{K}(\mathbf{h}; \nu_{12}, a_{12}) \\ \rho_{12} \mathbf{K}(\mathbf{h}; \nu_{12}, a_{12}) & \mathbf{K}(\mathbf{h}; \nu_2, a_2) \end{pmatrix} \begin{pmatrix} \sigma_1 \mathbf{P}_{\mathbf{t}}^T \\ \sigma_2 \mathbf{Q}_{\mathbf{t}}^T \end{pmatrix}, \end{aligned}$$

where $\nu_1, \nu_2 > 1$.

S.1.5 Proof of Proposition 1

The tangential vector field $\mathbf{Y}_{\text{tan},\mathbf{Z}}(\mathbf{s})$ with cross-covariance function $\mathbf{C}_{\text{tan},\mathbf{Z}}(\mathbf{s}, \mathbf{t})$ can be converted to the canonical coordinates $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ by the following transformation

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

where

$$\mathbf{T}_{\mathbf{s}} = \begin{pmatrix} -\sin \phi_{\mathbf{s}} & \cos \phi_{\mathbf{s}} & 0 \\ -\cos \theta_{\mathbf{s}} \cos \phi_{\mathbf{s}} & -\cos \theta_{\mathbf{s}} \sin \phi_{\mathbf{s}} & \sin \theta_{\mathbf{s}} \end{pmatrix}.$$

Then the cross-covariance function of \mathbf{V} is

$$\mathbf{C}_{\mathbf{V}}(\mathbf{s}, \mathbf{t}) = \mathbf{T}_{\mathbf{s}} \mathbf{C}_{\text{tan},\mathbf{Z}}(\mathbf{s}, \mathbf{t}) \mathbf{T}_{\mathbf{t}}^T. \quad (\text{S.3})$$

Note that

$$\mathbf{T}_{\mathbf{s}} \mathbf{P}_{\mathbf{s}} = \begin{pmatrix} -\sin \phi_{\mathbf{s}} & \cos \phi_{\mathbf{s}} & 0 \\ -\cos \theta_{\mathbf{s}} \cos \phi_{\mathbf{s}} & -\cos \theta_{\mathbf{s}} \sin \phi_{\mathbf{s}} & \sin \theta_{\mathbf{s}} \end{pmatrix},$$

and

$$\mathbf{T}_{\mathbf{s}} \mathbf{Q}_{\mathbf{s}} = \begin{pmatrix} \cos \theta_{\mathbf{s}} \cos \phi_{\mathbf{s}} & \cos \theta_{\mathbf{s}} \sin \phi_{\mathbf{s}} & -\sin \theta_{\mathbf{s}} \\ -\sin \phi_{\mathbf{s}} & \cos \phi_{\mathbf{s}} & 0 \end{pmatrix}.$$

Plugging them into (S.3) with $\mathbf{s} = \mathbf{t}$, we have

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

which is constant with respect to \mathbf{s} .

We can intuitively understand that the co-located cross-correlation between u and v has to be zero through computing it based on an alternative representation of \mathbf{V} , i.e., (8) and (9). Since Z_1 is isotropic, its change rates along latitudinal and longitudinal directions (at

the same location) are uncorrelated ([Adler and Taylor, 2009](#), page 116, (5.7.3)), i.e.,

$$\text{Cov} \left(\frac{\partial Z_1(\theta, \phi)}{\partial \phi}, \frac{\partial Z_1(\theta, \phi)}{\partial \theta} \right) = 0.$$

Similarly, the same holds for Z_2 . Moreover, the cross-covariance between Z_1 and Z_2 is also isotropic. Due to the curvature of the sphere, the covariance between the change rates of Z_1 and Z_2 along latitudinal direction (at the same location) equals the counterpart along longitudinal direction multiplied by $\sin^2 \theta$, i.e.,

$$\text{Cov} \left(\frac{\partial Z_1(\theta, \phi)}{\partial \phi}, \frac{\partial Z_2(\theta, \phi)}{\partial \phi} \right) = \text{Cov} \left(\frac{\partial Z_2(\theta, \phi)}{\partial \theta}, \frac{\partial Z_1(\theta, \phi)}{\partial \theta} \right) \sin^2 \theta.$$

Using the above two facts and after certain algebraic manipulations, we arrive at the same conclusion.

S.1.6 Proof of Proposition 2

According to the construction of the TMM, for any (θ, ϕ) except at the two poles,

$$\begin{aligned} u(\theta, \phi) \hat{\mathbf{u}} + v(\theta, \phi) \hat{\mathbf{v}} &= \nabla_{\mathbf{s}}^* \tilde{Z}_1(\mathbf{s}) + L_{\mathbf{s}}^* \tilde{Z}_2(\mathbf{s}) \quad \mathbb{P}\text{-a.e.} \\ &= \left(\frac{1}{\sin \theta} \frac{\partial \tilde{Z}_1}{\partial \phi} + \frac{\partial \tilde{Z}_2}{\partial \theta} \right) \hat{\phi} - \left(\frac{1}{\sin \theta} \frac{\partial \tilde{Z}_2}{\partial \phi} - \frac{\partial \tilde{Z}_1}{\partial \theta} \right) \hat{\theta} \\ &= \left(\frac{1}{\sin \theta} \frac{\partial \tilde{Z}_1}{\partial \phi} + \frac{\partial \tilde{Z}_2}{\partial \theta} \right) \hat{\mathbf{u}} + \left(\frac{1}{\sin \theta} \frac{\partial \tilde{Z}_2}{\partial \phi} - \frac{\partial \tilde{Z}_1}{\partial \theta} \right) \hat{\mathbf{v}}. \end{aligned}$$

Thus,

$$u(\theta, \phi) = \frac{1}{\sin \theta} \frac{\partial \tilde{Z}_1}{\partial \phi} + \frac{\partial \tilde{Z}_2}{\partial \theta} \quad \mathbb{P}\text{-a.e.},$$

and

$$v(\theta, \phi) = \frac{1}{\sin \theta} \frac{\partial \tilde{Z}_2}{\partial \phi} - \frac{\partial \tilde{Z}_1}{\partial \theta} \quad \mathbb{P}\text{-a.e.}$$

By condition **A1** and the chain rule for partial derivatives in quadratic mean (Potthoff, 2010, Theorem 2.13), (8) and (9) follow.

We have assumed that the cross-covariance function of $(Z_1(\mathbf{s}), Z_2(\mathbf{s}))^T$ (restricted to \mathbb{S}^2) only depends on the chordal distance, which is a function of θ_1, θ_2 and $\phi_1 - \phi_2$. Moreover, the coefficients of the partial derivatives are independent of longitude. Using the same argument as Jun and Stein (2008), we conclude that u and v are axially symmetric both marginally and jointly.

S.1.7 Implementation of the DFT for the Cross-covariance Matrix

Suppose that the observations are on a regular latitude-longitude grid on the unit sphere $\{(\theta_i, \phi_j) : i = 1, \dots, n_{\text{lat}}, j = 1, \dots, n_{\text{lon}}\}$, where θ and ϕ represent the co-latitude and longitude, respectively. Besides, $\theta_i - \theta_{i-1} = (\theta_{n_{\text{lat}}} - \theta_1)/(n_{\text{lat}} - 1)$ and $\phi_j - \phi_{j-1} = 2\pi/n_{\text{lon}}$, $\phi_1 = 0$, $\phi_{n_{\text{lon}}} = 2\pi - 2\pi/n_{\text{lon}}$. Let $(Y_1(\theta_i, \phi_j), Y_2(\theta_i, \phi_j))^T, i = 1, \dots, n_{\text{lat}}, j = 1, \dots, n_{\text{lon}}$ denote the observations. We rearrange them into $(\mathbf{Y}(\theta_1)^T, \dots, \mathbf{Y}(\theta_{n_{\text{lat}}})^T)^T$, where

$$\mathbf{Y}(\theta_i) = (\mathbf{Y}_1(\theta_i)^T, \mathbf{Y}_2(\theta_i)^T)^T = (Y_1(\theta_i, \phi_1), \dots, Y_1(\theta_i, \phi_{n_{\text{lon}}}), Y_2(\theta_i, \phi_1), \dots, Y_2(\theta_i, \phi_{n_{\text{lon}}}))^T.$$

Since the cross-covariance function of the random field \mathbf{Y} is axially symmetric,

$$\begin{aligned} \text{Cov}(\mathbf{Y}(\theta_i), \mathbf{Y}(\theta_j)) &= \mathbb{E}(\mathbf{Y}(\theta_i)\mathbf{Y}(\theta_j)^T) \\ &= \begin{pmatrix} \mathbb{E}(\mathbf{Y}_1(\theta_i)\mathbf{Y}_1(\theta_j)^T) & \mathbb{E}(\mathbf{Y}_1(\theta_i)\mathbf{Y}_2(\theta_j)^T) \\ \mathbb{E}(\mathbf{Y}_2(\theta_i)\mathbf{Y}_1(\theta_j)^T) & \mathbb{E}(\mathbf{Y}_2(\theta_i)\mathbf{Y}_2(\theta_j)^T) \end{pmatrix} \\ &= \begin{pmatrix} [C_{11}(\theta_i, \theta_j, \phi_k - \phi_l)]_{1 \leq k, l \leq n_{\text{lon}}} & [C_{12}(\theta_i, \theta_j, \phi_k - \phi_l)]_{1 \leq k, l \leq n_{\text{lon}}} \\ [C_{21}(\theta_i, \theta_j, \phi_k - \phi_l)]_{1 \leq k, l \leq n_{\text{lon}}} & [C_{22}(\theta_i, \theta_j, \phi_k - \phi_l)]_{1 \leq k, l \leq n_{\text{lon}}} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{C}_{11}(\theta_i, \theta_j) & \mathbf{C}_{12}(\theta_i, \theta_j) \\ \mathbf{C}_{21}(\theta_i, \theta_j) & \mathbf{C}_{22}(\theta_i, \theta_j) \end{pmatrix}, \end{aligned}$$

where $\mathbf{C}_{..}(\theta_i, \theta_j)$ are circulant matrices of the form (omitting the subscripts $..$)

$$\begin{pmatrix} c_0 & c_{n_{\text{lon}}-1} & \cdots & c_1 \\ c_1 & c_0 & \cdots & c_2 \\ \vdots & \vdots & & \vdots \\ c_{n_{\text{lon}}-1} & c_{n_{\text{lon}}-2} & \cdots & c_0 \end{pmatrix}.$$

A circulant matrix can be diagonalized by the DFT matrix \mathbf{F} , i.e.,

$$\mathbf{F}\mathbf{C}_{..}(\theta_i, \theta_j)\mathbf{F}^{-1} = \mathbf{\Lambda}_{..},$$

where

$$\mathbf{F} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \omega^1 & \cdots & \omega^{n_{\text{lon}}-1} \\ 1 & \omega^2 & \cdots & \omega^{2(n_{\text{lon}}-1)} \\ \vdots & \vdots & & \vdots \\ 1 & \omega^{n_{\text{lon}}-1} & \cdots & \omega^{(n_{\text{lon}}-1)(n_{\text{lon}}-1)} \end{pmatrix},$$

and $\omega = \exp(-2\pi i/n_{\text{lon}})$. Besides, the diagonal matrix $\mathbf{\Lambda}_{..}$ can be obtained simply by

$$\mathbf{\Lambda}_{..} = \text{diag}(\mathbf{F}\mathbf{c}_{..}),$$

where $\mathbf{c}_{..} = (c_0, \cdots, c_{n_{\text{lon}}-1})^T$. Thus, we transform the observations as follows

$$\begin{aligned} \mathbf{Y}(\theta_i) &= (\mathbf{Y}_1(\theta_i)^T, \mathbf{Y}_2(\theta_i)^T)^T \longrightarrow \\ \left(\tilde{\mathbf{Y}}_1(\theta_i)^T, \tilde{\mathbf{Y}}_2(\theta_i)^T \right)^T &= \left((\mathbf{F}\mathbf{Y}_1(\theta_i))^T / \sqrt{n_{\text{lon}}}, (\mathbf{F}\mathbf{Y}_2(\theta_i))^T / \sqrt{n_{\text{lon}}} \right)^T, \end{aligned}$$

so that

$$\begin{aligned}
\mathbb{E} \left(\tilde{\mathbf{Y}}_{\cdot}(\theta_i) \tilde{\mathbf{Y}}_{\cdot}(\theta_j)^{\text{H}} \right) &= \mathbf{F} \mathbb{E} \left(\mathbf{Y}_{\cdot}(\theta_i) \mathbf{Y}_{\cdot}(\theta_j)^{\text{T}} \right) \mathbf{F}^{\text{H}} / n_{\text{lon}} \\
&= \mathbf{F} \mathbb{E} \left(\mathbf{Y}_{\cdot}(\theta_i) \mathbf{Y}_{\cdot}(\theta_j)^{\text{T}} \right) \mathbf{F}^{-1} \\
&= \mathbf{\Lambda}_{\cdot\cdot},
\end{aligned}$$

where \cdot^{H} represents the Hermitian transpose. We further rearrange the transformed observations into

$$\begin{aligned}
&\left(\tilde{Y}_1(\theta_1, \phi_1), \dots, \tilde{Y}_1(\theta_{n_{\text{lat}}}, \phi_1), \dots, \tilde{Y}_1(\theta_1, \phi_{n_{\text{lon}}}), \dots, \tilde{Y}_1(\theta_{n_{\text{lat}}}, \phi_{n_{\text{lon}}}), \right. \\
&\quad \left. \tilde{Y}_2(\theta_1, \phi_1), \dots, \tilde{Y}_2(\theta_{n_{\text{lat}}}, \phi_1), \dots, \tilde{Y}_2(\theta_1, \phi_{n_{\text{lon}}}), \dots, \tilde{Y}_2(\theta_{n_{\text{lat}}}, \phi_{n_{\text{lon}}}) \right)^{\text{T}}.
\end{aligned}$$

Then the corresponding cross-covariance matrix has the form

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_1 & \mathbf{\Sigma}_{12} \\ \mathbf{\Sigma}_{12}^{\text{H}} & \mathbf{\Sigma}_2 \end{pmatrix},$$

where $\mathbf{\Sigma}_1$, $\mathbf{\Sigma}_{12}$ and $\mathbf{\Sigma}_2$ are complex block diagonal matrices. We know that the determinant of $\mathbf{\Sigma}$

$$\det(\mathbf{\Sigma}) = \det(\mathbf{\Sigma}_1 - \mathbf{\Sigma}_{12} \mathbf{\Sigma}_2^{-1} \mathbf{\Sigma}_{12}^{\text{H}}) \det(\mathbf{\Sigma}_2),$$

and the inverse of $\mathbf{\Sigma}$

$$\mathbf{\Sigma}^{-1} = \begin{pmatrix} (\mathbf{\Sigma}_1 - \mathbf{\Sigma}_{12} \mathbf{\Sigma}_2^{-1} \mathbf{\Sigma}_{12}^{\text{H}})^{-1} & -\mathbf{\Sigma}_1^{-1} \mathbf{\Sigma}_{12} (\mathbf{\Sigma}_2 - \mathbf{\Sigma}_{12}^{\text{H}} \mathbf{\Sigma}_1^{-1} \mathbf{\Sigma}_{12})^{-1} \\ (-\mathbf{\Sigma}_1^{-1} \mathbf{\Sigma}_{12} (\mathbf{\Sigma}_2 - \mathbf{\Sigma}_{12}^{\text{H}} \mathbf{\Sigma}_1^{-1} \mathbf{\Sigma}_{12})^{-1})^{\text{H}} & (\mathbf{\Sigma}_2 - \mathbf{\Sigma}_{12}^{\text{H}} \mathbf{\Sigma}_1^{-1} \mathbf{\Sigma}_{12})^{-1} \end{pmatrix}.$$

Thus, both $\det(\mathbf{\Sigma})$ and $\mathbf{\Sigma}^{-1}$ can be computed efficiently.

S.1.8 Initial Value Specification

To specify initial values for the model parameters, we first fix the spatial scale parameter at a sufficiently large value (e.g., 5 in the simulation study and 10 in the real data analysis).

This is motivated by the numerical results in [Kaufman and Shaby \(2013\)](#) that when the spatial scale parameter is fixed at some value larger than its true value, both the estimators and predictors can still perform well. We set initial values for τ_1 and τ_2 at certain relatively small values since the specification almost does not affect the parameter estimation. For the remaining parameters, we randomly select 100 parameter vectors in the parameter space using Latin hypercube sampling (LHS), and choose the one with the smallest negative log-likelihood.

S.1.9 Computational Details of Fitting the NBG

This subsection describes the computational details of fitting the NBG. Note that currently there is no publicly available package for the NBG. Since the NGB consists of two components: the linear combinations of partial derivatives of a scalar random field and the PARS-BM with $\rho_{12} = 0$, one can start with estimating the parameters in the second component by ignoring the first one (i.e., assuming $a_i = b_i = 0$), and then conditionally on the parameter estimates in the second component, estimate the remaining parameters in the first component. We use the two-step parameter estimates as initial values in the joint minimization of the negative log-likelihood function. The *interior-point* algorithm through the Matlab function *fmincon* is used to jointly minimize the negative log-likelihood function in all cases except computing bootstrap standard errors. The maximum number of function evaluations for the algorithm is specified as 20,000 to estimate the parameters in the real data analysis, and 10,000 otherwise. The computation of bootstrap standard errors is time-consuming since the estimation procedure needs to be repeated 200 times. To reduce the computational burden with a mild loss of accuracy, we use the *Nelder-Mead simplex direct search* algorithm through the Matlab function *fminsearch* in the estimation procedure. Although this is an unconstrained optimization method, the results are not significantly affected since initial values can be specified close to the actual minimizer.

S.2 Additional Numerical Results

S.2.1 Accuracy of Parameter Estimation on Irregular Grids

To investigate the accuracy of parameter estimation on irregular grids, we conduct some smaller-scale experiments (due to the computational burden). Specifically, we generate 100 realizations with

$$\boldsymbol{\theta} = (1, 1, 0.5, 3, 4, 1/2, 0.1, 0.1)$$

on two HEALPix grids (Górski et al., 2005) with 192 and 768 grid points, respectively. The HEALPix grid partitions the unit sphere into equal area pixels, and thus is naturally irregular.

The results are summarized in Figure S.1.

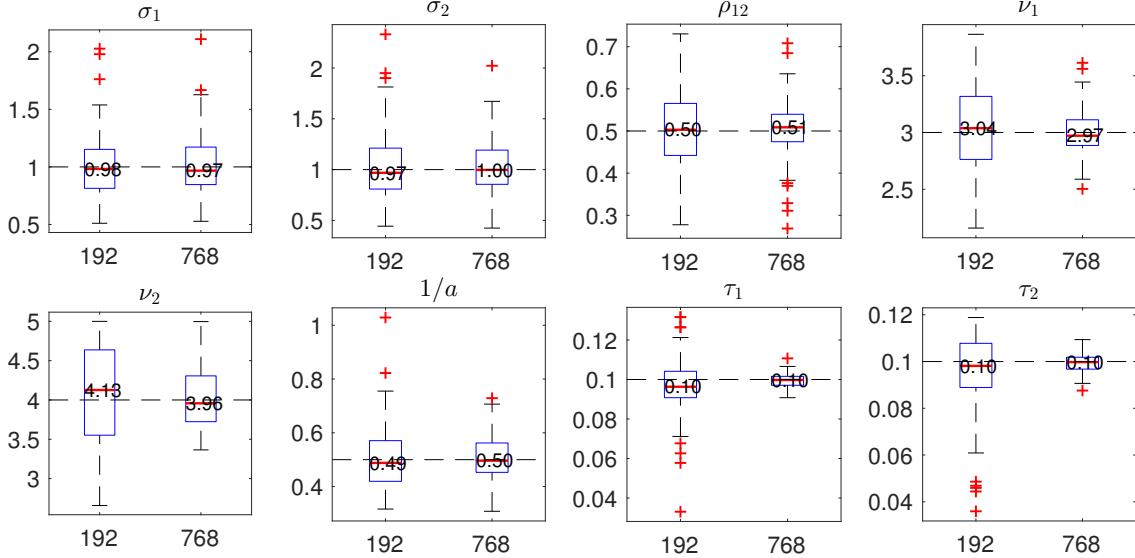


Figure S.1: Results of the Monte Carlo simulation study to investigate the accuracy of parameter estimation for the TMM on irregular grids. The MLEs of $\boldsymbol{\theta} = (\sigma_1, \sigma_2, \rho_{12}, \nu_1, \nu_2, 1/a, \tau_1, \tau_2)$ are summarized by boxplots for two simulated data sets with increasing sample size (shown on x-axis). The dashed horizontal lines are at the true values.

S.2.2 Accuracy of Parameter Estimation For Three Different Noise Levels

It is also interesting to see how the noise level affects the accuracy of parameter estimation.

Specifically, we consider three different noise levels, and generate 500 realizations with

$$\boldsymbol{\theta} = (1, 1, 0.5, 3, 4, 1/2, 0.1, 0.1) \quad \text{low noise}$$

$$\boldsymbol{\theta} = (1, 1, 0.5, 3, 4, 1/2, 0.2, 0.2) \quad \text{medium noise}$$

$$\boldsymbol{\theta} = (1, 1, 0.5, 3, 4, 1/2, 0.3, 0.3) \quad \text{high noise}$$

on a regular grid with $(n_{\text{lat}}, n_{\text{lon}}) = (15, 30)$, respectively. Figure S.2 shows that the higher the noise level, the larger the standard errors of the estimates, especially for ν_1 and ν_2 .

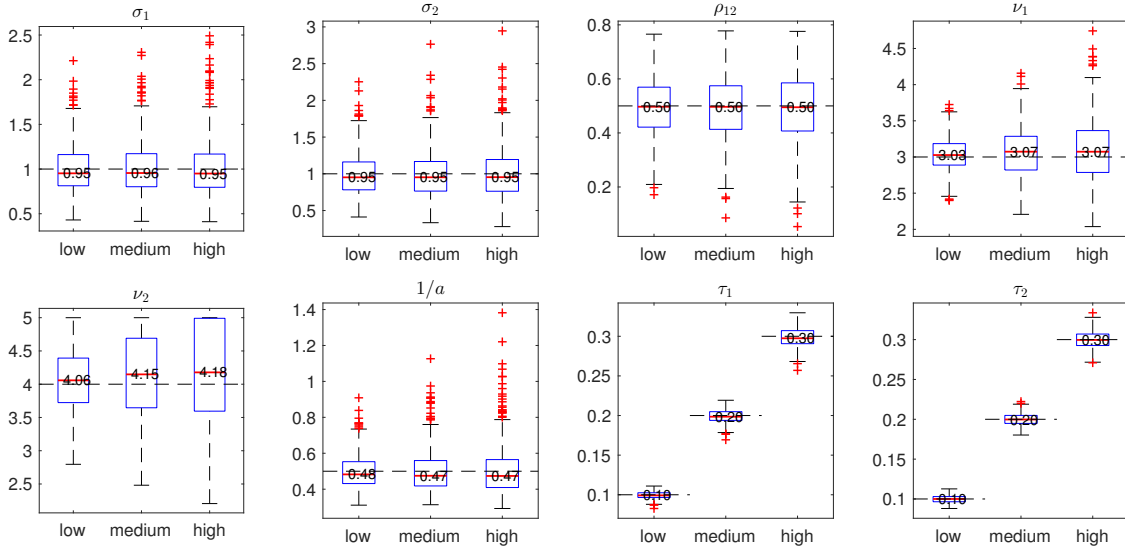


Figure S.2: Results of the Monte Carlo simulation study to investigate the accuracy of parameter estimation for the TMM for three different noise levels (low: $\tau_1 = \tau_2 = 0.1$; medium: $\tau_1 = \tau_2 = 0.2$; high: $\tau_1 = \tau_2 = 0.3$). The MLEs of $\boldsymbol{\theta} = (\sigma_1, \sigma_2, \rho_{12}, \nu_1, \nu_2, 1/a, \tau_1, \tau_2)$ are summarized by boxplots. The dashed horizontal lines are at the true values.

S.2.3 Accuracy of Parameter Estimation For Two Different Sets of Smoothness Parameters

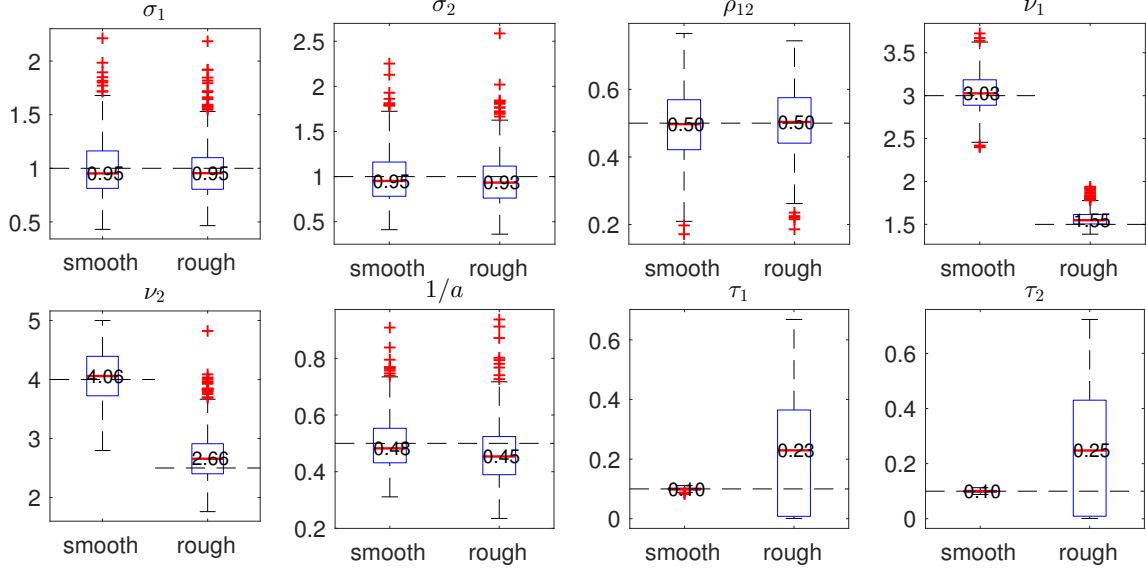


Figure S.3: Results of the Monte Carlo simulation study to investigate the accuracy of parameter estimation for the TMM for two different sets of smoothness parameters (smooth: $\nu_1 = 3, \nu_2 = 4$; rough: $\nu_1 = 1.5, \nu_2 = 2.5$). The MLEs of $\theta = (\sigma_1, \sigma_2, \rho_{12}, \nu_1, \nu_2, 1/a, \tau_1, \tau_2)$ are summarized by boxplots. The dashed horizontal lines are at the true values.

We also test the effect of the smoothness of the field on the accuracy of parameter estimation. Specifically, we consider two different sets of smoothness parameters, and generate 500 realizations with

$$\theta = (1, 1, 0.5, 3, 4, 1/2, 0.1, 0.1) \quad \text{smooth}$$

$$\theta = (1, 1, 0.5, 1.5, 2.5, 1/2, 0.1, 0.1) \quad \text{rough}$$

on a regular grid with $(n_{\text{lat}}, n_{\text{lon}}) = (15, 30)$, respectively.

The results are summarized in Figure S.3. Due to the differential operators that have been applied, the latter specification yields a field effectively as rough as the one generated by the Matérn model with $\nu = 1.5 - 1 = 0.5$. We know that it is difficult to distinguish between a field and observational (white) noise when the former is very rough. Thus, in the

second case, τ_1 and τ_2 are significantly overestimated, and the estimates have large variability. Moreover, compared with the first case, the upward biases of the estimates for ν_1 and ν_2 are more pronounced.

S.2.4 Accuracy of Parameter Estimation When Covariates Are Included in the Model

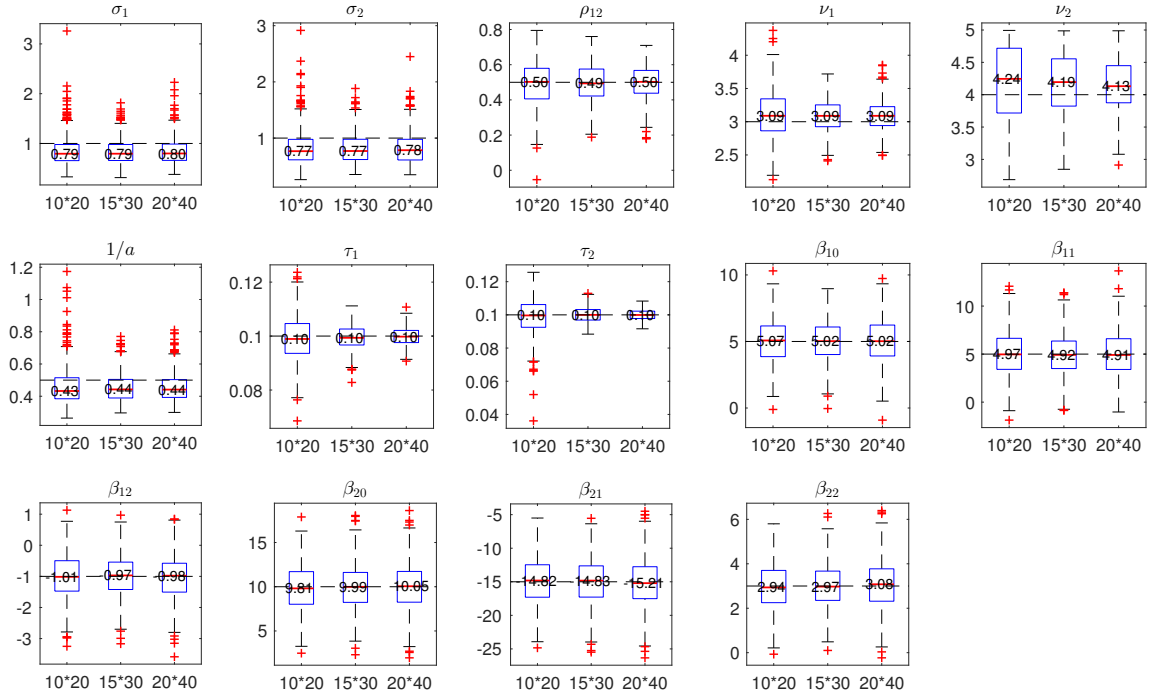


Figure S.4: Results of the Monte Carlo simulation study to investigate the accuracy of parameter estimation for the TMM when covariates are included in the model. The MLEs of $\boldsymbol{\theta} = (\sigma_1, \sigma_2, \rho_{12}, \nu_1, \nu_2, 1/a, \tau_1, \tau_2)$ and $\boldsymbol{\beta}_i, i = 0, 1, 2$ 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.

We include a polynomial function of the co-latitude as covariates in the model such that for the field $\mathbf{Y}(\mathbf{s}), \mathbf{s} = (\theta, \phi)$, its mean is $\beta_0 + \beta_1\theta + \beta_2\theta^2$, where $\boldsymbol{\beta}_i = (\beta_{1i}, \beta_{2i})^T, i = 0, 1, 2$ are two-dimensional coefficient vectors, and its cross-covariance function still follows the TMM

(augmented with nugget effects). We generate 500 realizations with

$$\boldsymbol{\theta} = (1, 1, 0.5, 3, 4, 1/2, 0.1, 0.1),$$

and

$$\boldsymbol{\beta}_0 = (5, 10)^T, \boldsymbol{\beta}_1 = (5, -15)^T, \boldsymbol{\beta}_2 = (-1, 3)^T$$

on three regular grids with $(n_{\text{lat}}, n_{\text{lon}}) = (10, 20), (15, 30)$ and $(20, 40)$, respectively. Figure S.4 shows that the estimates for σ_1, σ_2 and $1/a$ have larger biases than those obtained without any covariate. This suggests difficulty in jointly estimating the parameters contained in both the mean and the random components. The biases of the estimates for $\boldsymbol{\beta}_i, i = 0, 1, 2$ are small, while the standard errors of the estimates do not decrease significantly as the sample size increases.

S.2.5 Bootstrap Standard Errors

Asymptotic standard errors of MLEs based on the Fisher information matrix are not suitable in our context since their validity depends on the framework of increasing domain asymptotics. Therefore, we seek to estimate the standard errors of MLEs using a parametric bootstrap procedure. Accordingly, the bootstrap samples are independent realizations of the TMM with its parameters estimated by the method of maximum likelihood (Horowitz, 2001). To test the effectiveness of the parametric bootstrap, using the simulation study in Section 3.2 with $(n_{\text{lat}}, n_{\text{lon}}) = (15, 30)$ as an example, we compare the empirical standard errors, and the bootstrap standard errors computed based on 200 bootstrap samples. The empirical standard errors are computed as the standard errors of the estimates obtained from the 500 simulation runs. We obtain 10 sets of bootstrap standard errors for the first 10 of the 500 simulation runs. Figure S.5 shows the ratios between the bootstrap and empirical standard errors. For each parameter, the 10 ratios are summarized by a boxplot. It is noticeable that the bootstrap standard errors tend to be slightly higher than the empirical ones, which

suggests that the former can be seen as conservative estimates for the true standard errors. Also, the ratios for σ_1 and σ_2 have relatively large spreads.

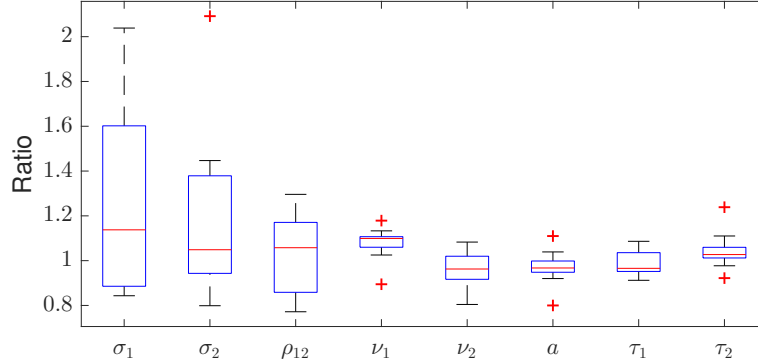


Figure S.5: Ratios between the bootstrap and empirical standard errors. For each parameter, the 10 ratios are summarized by a boxplot.

S.2.6 Spatial Prediction

In this subsection, we compare the predictive performance of the TMM with the parsimonious bivariate Matérn model (PARS-BM) (see Section ??) by simulation. The spatial prediction of vector fields can be achieved by cokriging; see Myers (1982); Ver Hoef and Cressie (1993) for examples. We use the implementation of the PARS-BM in the R package RandomFields (Version 3.0.62) (Schlather et al., 2015). Suppose that the data are simulated from the TMM on the HEALPix grid with 768 grid points using the same parameter specification as (?). Half of the locations are randomly selected outside a randomly selected longitudinal region with width 30° for estimation, and the remaining locations are held out for prediction. In this way, both short-range and long-range predictions are taken into consideration. We assess the prediction accuracy using several 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). The cross-validation procedure is repeated 500 times, and the boxplots of the four scoring rules for the two models are shown in Figure S.6. All the boxplots of the TMM are lower than those of the PARS-BM, which indicates that the TMM has better predictive performance in terms of the four scoring rules.

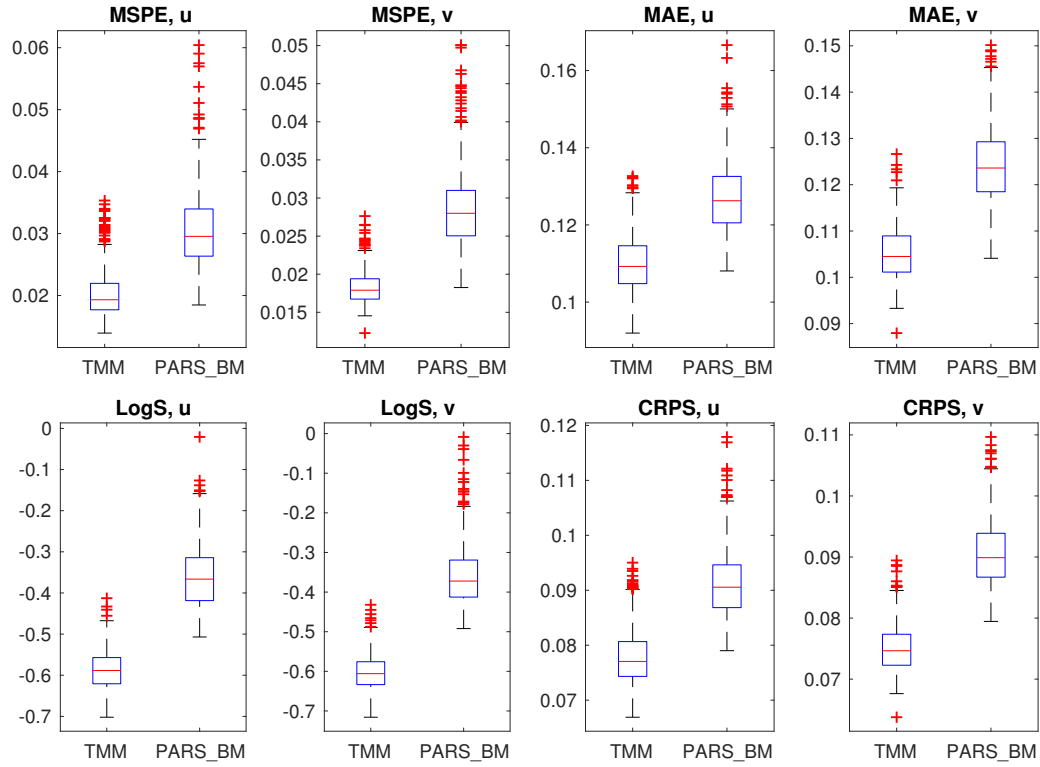


Figure S.6: Boxplots of the four scoring rules MSPE, MAE, LogS, and CRPS for the TMM and the PARS-BM with 500 replications.

S.3 Additional Data Example Results

S.3.1 Temporal Independence and Normality Check

Figures S.7 and S.8 display the sample autocorrelation functions of the u and v residual wind fields, respectively, at 16 randomly selected locations in the subregion of the Indian Ocean. They support our assumption that the residuals are temporally uncorrelated.

The marginal Q-Q plots and the Chi-Square Q-Q plot, shown in Figure S.9 for January, 2000, suggest that the Gaussianity assumption on the u and v residual wind fields is reasonable.

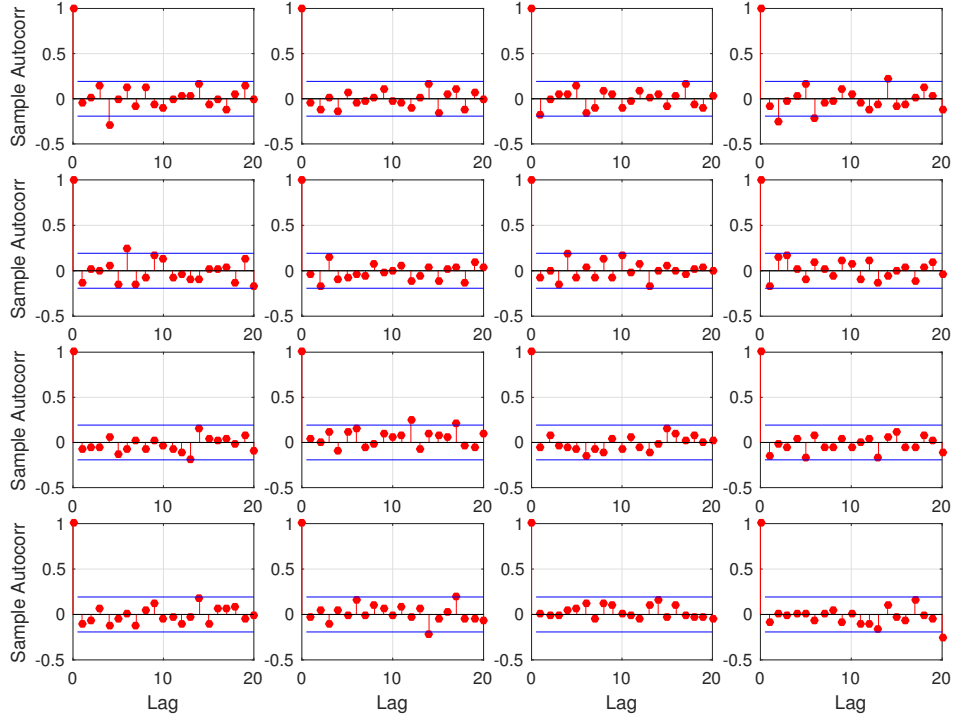


Figure S.7: Sample autocorrelation functions of the u residual wind fields at 16 randomly selected locations in the subregion of the Indian Ocean.

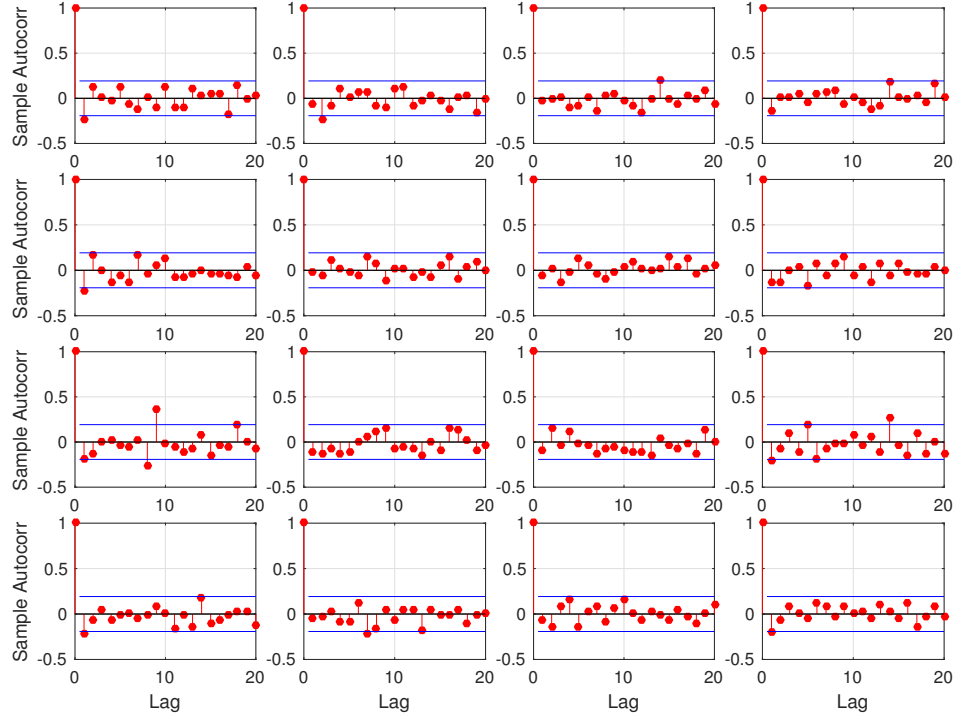


Figure S.8: Sample autocorrelation functions of the v residual wind fields at 16 randomly selected locations in the subregion of the Indian Ocean.

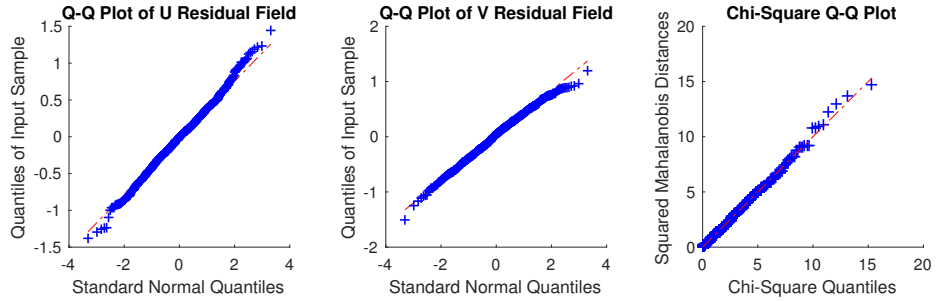


Figure S.9: Marginal Q-Q plots (the first two) and Chi-Square Q-Q plot (the third) of the u and v residual wind fields for January, 2000 in the subregion of the Indian Ocean.

S.3.2 Comparison between the Empirical and Fitted Variances of the u and v Residual Fields

Figure S.10 displays the empirical and fitted variances of the u and v residual wind fields under the TMM, PARS-BM and NBG, which are plotted with respect to latitude and longitude. The fitted variances of the u and v residual wind fields under the TMM and PARS-BM are constant due to Proposition 1 and the isotropy of the model, respectively. Although the NBG is supposed to be non-stationary with respect to latitude, the trend slopes of the fitted variances plotted with respect to latitude are negligible. In general, the fits of the three models are close to each other. However, there is a noticeable discrepancy between the empirical and fitted variances, especially for the u residual wind field, whose empirical variances have a pattern depending on latitude. This is not surprising because all the three models are too simple to capture the displayed patterns given the number of parameters they have. As discussed in Section 5, the discrepancy for the TMM can be potentially alleviated by assuming a more complicated underlying potential field, such as an anisotropic one.

S.3.3 Two Sample T-Tests

Table S.1: P-values of the two sample t -tests comparing the model pairs (TMM, NMG) and (TMM, PARS-BM)

| Models | Variable | MSPE | MAE | LogS | CRPS |
|---------------|----------|------------|------------|------------|------------|
| TMM & NMG | u | 0.0085 | 0.0291 | 0.0304 | 0.0212 |
| | v | 0.0882 | 0.4628 | 0.7372 | 0.3637 |
| TMM & PARS-BM | u | 1.3682e-12 | 2.5031e-11 | 8.5639e-11 | 8.7630e-12 |
| | v | 3.1706e-06 | 2.1925e-07 | 7.1738e-05 | 1.5008e-06 |

Table S.1 displays the p -values of the two sample t -tests, which compare the model pairs (TMM, NMG) and (TMM, PARS-BM). Since the comparison of each model pair involves

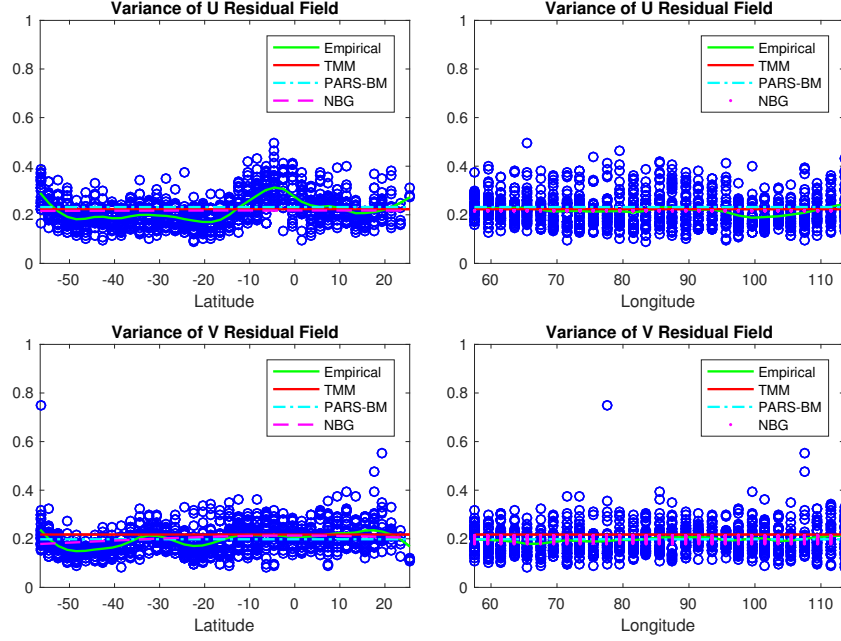


Figure S.10: Empirical and fitted variances of the u and v residual fields: circles for the empirical variances, green solid lines for the loess curves fitted to the empirical variances, red solid lines for the TMM, dash-dot lines for the PARS-BM, and dashed lines (left) or dots (right) for the NBG. They are plotted with respect to latitude (left) and longitude (right).

multiple statistical tests, we use the Bonferroni correction to adjust the level of significance α , and hence it becomes $\alpha/n = 0.05/8 = 0.00625$. Under this threshold, none of the t-tests for the TMM and NMG is statistically significant, while all of those for the TMM and PARS-BM are significant.

S.3.4 Predicted and Observed Wind Field Comparison

In this subsection, we compare the wind fields predicted by the TMM, PARS-BM and NBG with the observations. The predicted wind fields are obtained by cokriging, which predicts at the locations in the rectangular region based on the observations outside the region, using the parameters estimated from the entire data set (see Table 1). Figure S.11 displays the comparison among the predicted and observed wind fields for January, 2000 in the rectangular region, together with a decomposition of the predicted wind field by the TMM into its curl-free and divergence-free components. Note that there are certain discrepancies

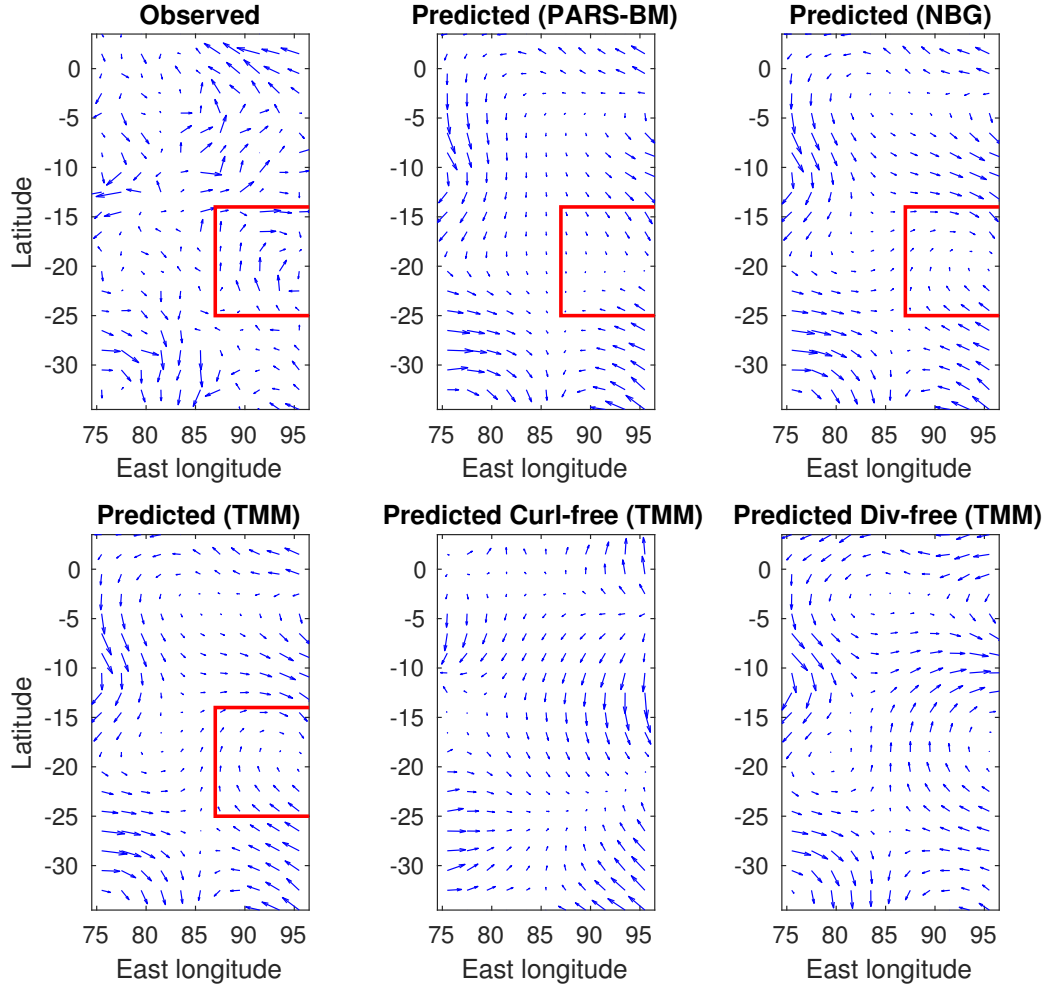


Figure S.11: Observed and predicted wind fields by the TMM, PARS-BM and NBG for January, 2000 in the rectangular region in the center of the subregion of the Indian Ocean. The predicted wind field by the TMM is decomposed into its curl-free and divergence-free components. Note that the arrows have been automatically scaled to fit within the grid.

between the predicted and observed wind fields, which are commonly seen in practice due to the errors resulted from long-distance predictions, and the inevitable inconsistency between statistical models and the reality. Nonetheless, it is worth pointing out that the predicted wind fields by the TMM and NBG capture certain fine details of the observations, which are missed by the PARS-BM. For example, in the marked red-color rectangle, the winds predicted by the TMM and NBG and the observed winds both rotate clockwise and show similar structures.

References

- Adler, R. J. and Taylor, J. E. (2009). *Random fields and geometry*. Springer Science & Business Media.
- 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.
- 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.
- Handcock, M. S. and Wallis, J. R. (1994). An approach to statistical spatial-temporal modeling of meteorological fields. *Journal of the American Statistical Association*, 89(426):368–378.
- Horowitz, J. L. (2001). The bootstrap. *Handbook of econometrics*, 5:3159–3228.
- 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.
- Kaufman, C. G. and Shaby, B. A. (2013). The role of the range parameter for estimation and prediction in geostatistics. *Biometrika*, 100(2):473–484.
- Myers, D. E. (1982). Matrix formulation of co-kriging. *Journal of the International Association for Mathematical Geology*, 14(3):249–257.
- Potthoff, J. (2010). Sample properties of random fields III: Differentiability. *Communications on Stochastic Analysis*, 4(3):335–353.

- 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.
- Ver Hoef, J. M. and Cressie, N. (1993). Multivariable spatial prediction. *Mathematical Geology*, 25(2):219–240.