Statistical Science
2017, Vol. 32, No. 4, 501–513
DOI: 10.1214/17-STS620

Spherical Process Models for Global Spatial Statistics

Jaehong Jeong, Mikyoung Jun and Marc G. Genton

Abstract. Statistical models used in geophysical, environmental, and climate science applications must reflect the curvature of the spatial domain in global data. Over the past few decades, statisticians have developed covariance models that capture the spatial and temporal behavior of these global data sets. Though the geodesic distance is the most natural metric for measuring distance on the surface of a sphere, mathematical limitations have compelled statisticians to use the chordal distance to compute the covariance matrix in many applications instead, which may cause physically unrealistic distortions. Therefore, covariance functions directly defined on a sphere using the geodesic distance are needed. We discuss the issues that arise when dealing with spherical data sets on a global scale and provide references to recent literature. We review the current approaches to building process models on spheres, including the differential operator, the stochastic partial differential equation, the kernel convolution, and the deformation approaches. We illustrate realizations obtained from Gaussian processes with different covariance structures and the use of isotropic and nonstationary covariance models through deformations and geographical indicators for global surface temperature data. To assess the suitability of each method, we compare their log-likelihood values and prediction scores, and we end with a discussion of related research problems.

Key words and phrases: Axial symmetry, chordal distance, geodesic distance, nonstationarity, smoothness, sphere.

1. INTRODUCTION

Rapid developments in science and technology have facilitated the generation of global data sets from many geophysical applications (Cressie and Johannesson, 2008). When working on a global scale, it is essential to respect the curvature of the spatial domain, particularly, with respect to spatial dependence. Geodesic (great circle) distance is physically the most natural distance metric for constructing covariance models for processes on the surface of a sphere (Gneiting, 2013). However, the mathematical limitations associated with covariance functions on spheres have encouraged the development of approximations based on the chordal distance instead. As Yadrenko (1983) and Yaglom (1987) pointed out, isotropic covariance functions defined on Euclidean spaces can be used on spheres when coupled with the chordal distance. The Matérn covariance functions often work better with the chordal distance than the geodesic distance if the process is smooth (Jeong and Jun, 2015a, Guinness and Fuentes, 2016). However, unless the range of the spatial correlation is small relative to the curvature of the earth, this approach should be used with caution; the chosen distances may result in different parameter estimations of the covariance structures and a poor pre-

Jaehong Jeong is Postdoctoral Fellow, CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia (e-mail: jaehong.jeong@kaust.edu.sa). Mikyoung Jun is Associate Professor, Department of Statistics, Texas A&M University, College Station, Texas 77843-3143, USA (e-mail: mjun@stat.tamu.edu). Marc G. Genton is Distinguished Professor, CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia (e-mail: marc.genton@kaust.edu.sa).

diction performance at new locations (Banerjee, 2005, Jeong and Jun, 2015b, Porcu, Bevilacqua and Genton, 2016).

Various covariance functions have been derived from different construction approaches. Marinucci and Peccati (2011) characterized isotropic Gaussian processes on spheres in detail for cosmological applications. Gneiting (2013) discussed methods for generating valid covariance functions based on the geodesic distance. Ma (2012) and Du, Ma and Li (2013) characterized covariance and variogram functions for vector-valued processes in terms of Gegenbauer polynomials, while Heaton et al. (2014) defined covariance functions via kernel convolution. Porcu, Bevilacqua and Genton (2016) and Alegría et al. (2017b) offered characterizations of spatio-temporal covariance and cross-covariance models.

In this paper, we discuss the issues that arise when modeling spherical data, and we review the current geostatistical approaches with a focus on covariance functions for data on a global scale. The remainder of the paper is organized as follows: Section 2 provides the theoretical background for processes on the surface of a sphere. Section 3 addresses covariance models on spheres in general and Section 4 reviews constructive approaches to modeling. Section 5 illustrates the realizations obtained from processes with different covariance structures and the performance of covariance models on global surface temperature data. The paper ends with a discussion in Section 6.

2. THEORETICAL BACKGROUND FOR PROCESSES ON A SPHERE

2.1 Definitions and Notations

Consider $\mathcal{S}^2 = \{\mathbf{x} \in \mathbb{R}^3 : \|\mathbf{x}\| = R\}$, where $\|\cdot\|$ is the Euclidean norm with center zero $(0,0,0)$ and radius $R \in (0,\infty)$, a two-dimensional (2D) spherical surface embedded in a three-dimensional (3D) Euclidean space, $\mathbb{R}^3$. We can parameterize a point on $\mathcal{S}^2$ as (R, L, ℓ), where $L \in [-\pi/2, \pi/2]$ is the latitude and $\ell \in [-\pi, \pi]$ is the longitude. If we use the Cartesian coordinate system, the same point is denoted by (x, y, z), and $x = R\cos\ell\cos L$, $y = R\sin\ell\cos L$, and $z = R\sin L$.

We now define $Z(L,\ell)$ as a stochastic process on $\mathcal{S}^2$ at the latitude L and the longitude ℓ. The process is isotropic if the covariance of the process depends only on the distance between the two points on $\mathcal{S}^2$. If Z is Gaussian, then for $\mathbf{Z} = \{Z(L_1,\ell_1),\ldots,Z(L_n,\ell_n)\}^\top$, we have $\mathbf{Z} \sim N_n(\boldsymbol{\mu}, \boldsymbol{C})$ with mean vector $\boldsymbol{\mu} = \{\mu(L_1,\ell_1),\ldots,\mu(L_n,\ell_n)\}^\top$ and covariance matrix $\boldsymbol{C}$, where $C_{ij} = \operatorname{cov}\{Z(L_i,\ell_i), Z(L_j,\ell_j)\}$, $i,j = 1,\ldots,n$. The covariance matrix should be positive definite, that is, $\mathbf{a}^\top \boldsymbol{C}\mathbf{a} \geq 0$ for any nonzero vector $\mathbf{a} \in \mathbb{R}^n$, any spatial locations $(L_i,\ell_i)_{i=1}^n$, and any integer n. For a spatio-temporal random field, we can define $Z(L,\ell,t)$ similarly, where t is the index for time.

As mentioned previously, the geodesic distance is the most natural distance metric on $\mathcal{S}^2$. The central angle between two points, (L_1,ℓ_1) and (L_2,ℓ_2), on $\mathcal{S}^2$ is

$$\theta = \arccos\{\sin L_1 \sin L_2 + \cos L_1 \cos L_2 \cos(\ell_1 - \ell_2)\}. \tag{1}$$

Therefore, the geodesic distance between these two points is given by $d_G = R\theta$. We can define the chordal distance, which is the Euclidean distance in $\mathbb{R}^3$, as $d_C = 2R\sin(\theta/2)$. Figure 1 illustrates the difference between the geodesic distance and the chordal distance on a one-dimensional circle, and shows how the discrepancy between those two distances changes with the distance between (L_1,ℓ_1) and (L_2,ℓ_2).

2.2 Smoothness of Processes on a Sphere

As argued by Guinness and Fuentes (2016), the notion of a great circle is useful for studying the derivatives of random functions on $\mathcal{S}^2$. Assume that $Z(\mathbf{x}), \mathbf{x} \in \mathcal{S}^2$, is a random process with an isotropic covariance function C and the Hilbert space of linear combinations H_C of Z, with finite variance. Suppose X is the collection of all points along one great circle. Then, there is a distance-preserving mapping $\tilde{\phi} : X \to [0, 2\pi)$ that associates each point on the great circle with a unique angle. For some choice of $\tilde{\phi}$, define $Z_X\{\tilde{\phi}(\mathbf{x})\} = Z(\mathbf{x})$ as the restriction of Z on X. Then Z_X and Z are mean-square differentiable at $\mathbf{x}$ if $Z'_X\{\tilde{\phi}(\mathbf{x})\} = \lim_{h_n\to 0}[Z_X\{\tilde{\phi}(\mathbf{x}) + h_n\} - Z_X\{\tilde{\phi}(\mathbf{x})\}]/h_n$ exists in H_C and if $Z'_X\{\tilde{\phi}(\mathbf{x})\}$ exists for every X that contains $\mathbf{x}$.

Efforts have been made to introduce conditions and covariance functions that can flexibly specify the smoothness of a process on $\mathcal{S}^2$. Hitczenko and Stein (2012) gave conditions for mean-square differentiability when a covariance function is expressed in terms of its spherical harmonic representation. Lang and Schwab (2015) gave further results on Hölder continuity and differentiability of sample paths of the process. According to Gneiting (2013), the geodesic distance cannot be used for any differentiable Matérn field. Jeong and Jun (2015a) studied processes that allow differentiability at the origin, similar to the way

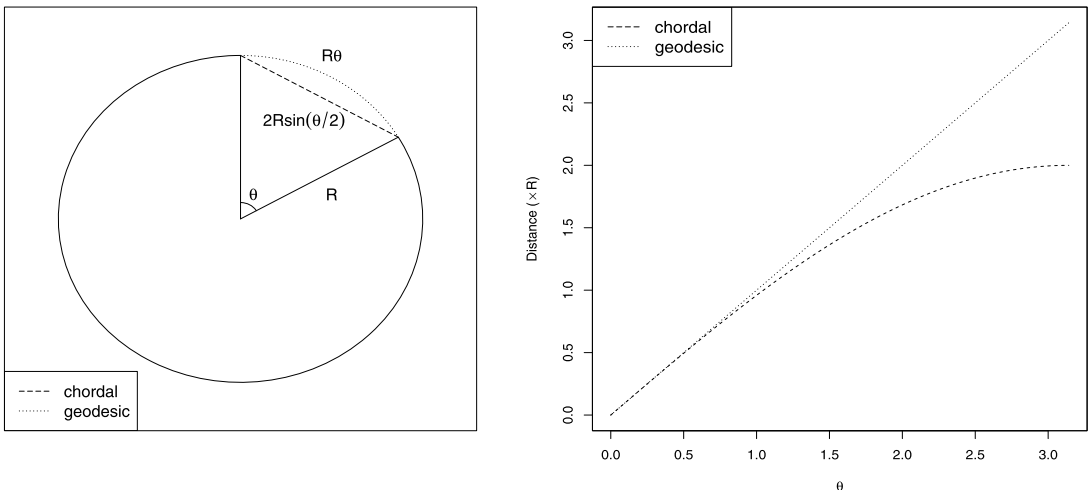

FIG. 1. *Chordal vs. geodesic distance on a circle with radius R.*

the Matérn covariance function does in Euclidean spaces. Guinness and Fuentes (2016) proposed a circular Matérn covariance function and provided conditions that can be used to determine the number of mean-square derivatives directly from the covariance function or from its Fourier series representation.

3. COVARIANCE MODELS ON A SPHERE

3.1 Validity of Covariance Functions on a Sphere

According to Schoenberg (1942) and Yaglom (1987), a covariance function of an isotropic process on $\mathcal{S}^2$ can be expressed as the infinite sum of Legendre polynomials, with nonnegative coefficients and the cosine of the geodesic distance as their arguments. Schreiner (1997) showed that such covariance functions are strictly positive definite if only finitely many coefficients are equal to zero. The challenge in covariance modeling for isotropic processes on $\mathcal{S}^2$ is that these infinite sums usually do not have closed forms. Even if there is a closed form, all of these infinite sums are analytic. A well-known example of such a closed form is the Poisson kernel:

$$K_r(\theta) = \frac{1-r^2}{(1-2r\cos\theta+r^2)^{3/2}}, \tag{2}$$
$$0 \le \theta \le \pi, 0 \le r < 1.$$

Das (2000) considered linear combinations of the Poisson kernel (or other lesser-known closed forms) and the truncated infinite summation of Legendre polynomials discussed above. This approach clearly provides valid covariance functions for isotropic processes on $\mathcal{S}^2$. However, all of the covariance functions obtained in this way are analytic, which may be too smooth for geophysical processes (Stein, 1999).

We could simply use a Euclidean distance-based isotropic covariance function on a plane for isotropic processes on $\mathcal{S}^2$. However, the resulting model may not be valid with geodesic distances. For example, the Gaussian covariance function, $C(\theta) = \exp(-\alpha\theta^2)$, $\alpha > 0$, is not positive definite on $\mathcal{S}^2$ with a geodesic distance unless $\alpha = 0$ (Gneiting, 1999a). Efforts have been made to examine the validity of several parametric covariance functions when used with the geodesic distance. Gneiting (1998) proposed methods for testing the validity of isotropic covariance functions on a circle, and Gneiting (1999b) provided conditions for the validity of a linear covariance function on d-dimensional balls. Huang, Zhang and Robeson (2011) and Gneiting (2013) examined the validity of the most popular covariance functions on $\mathcal{S}^2$, for example, the Matérn class, the spherical model, and the Cauchy covariance functions.

Yaglom (1987) provided a simple way of using isotropic covariance functions based on the Euclidean distance for processes on $\mathcal{S}^2$. The key is using the chordal distance rather than the geodesic distance on the sphere. This method can provide a rich class of isotropic covariance functions allowing the interpretation of scale, support, shape, smoothness, and fractal index for processes (Gneiting, 2013). However, the chordal distance approach results in limited flexibility by imposing a lower bound on the isotropic correlation function, rejecting correlation values below -0.218 from the integral representation of isotropic correlation functions in $\mathbb{R}^3$ (Stein, 1999), and making no allowances for the curvature of the earth. In fact, it is

locally linear and, according to Gneiting (2013), it is counter to spherical geometry for large geodesic distance values, and thus may bring about physically unrealistic distortions.

3.2 Isotropic Covariance Models on a Sphere

Ma (2015) extended the characterizations of isotropic covariance matrix functions on $\mathcal{S}^2$ via Gegenbauer polynomials from Ma (2012). For example, a function $C(\theta)$ whose entries are continuous on $[0, \pi]$ is a covariance function on all spheres ($\mathcal{S}^\infty$) if and only if $C(\theta) = \sum_{k=0}^{\infty} b_k (\cos\theta)^k$ or $C(\theta) = g(\cos\theta)$, where b_k is a summable sequence of nonnegative numbers, $g(\cdot)$ is continuous on $[-1, 1]$, and $g(x) + g(-x)$ and $g(x) - g(-x)$ are absolutely monotone on $[0, 1]$.

Porcu, Bevilacqua and Genton (2016) proposed stationary covariance functions for spatio-temporal processes on $\mathcal{S}^2$, as well as cross-covariance functions for multivariate random processes defined on $\mathcal{S}^2$. They showed that a class of spatio-temporal covariance functions based on the Euclidean distance (Gneiting, 2002) can be adapted for processes over time on $\mathcal{S}^2$. Alegría et al. (2017b) extended this to spatio-temporal cross-covariance functions on $\mathcal{S}^2$. Porcu, Bevilacqua and Genton (2016) also presented a direct construction of $C(\theta, u) = \sum_{k=0}^{\infty} g_k(u)(\cos\theta)^k$, $(\theta, u) \in [0, \pi] \times \mathbb{R}$, where $\{g_k(\cdot)\}_{k=0}^{\infty}$ is an absolutely convergent sequence of continuous and positive-definite functions.

Closed-form expressions for cross-covariance functions were also introduced by Porcu, Bevilacqua and Genton (2016) using the geodesic distance and basing conditions for a valid construction on a completely monotone function and a Bernstein function. Porcu and Schilling (2011) and Porcu and Zastavnyi (2011) gave examples of these two classes of functions. Porcu, Bevilacqua and Genton (2016) defined the mapping $\boldsymbol{C}(\theta) = \{\sigma_i \sigma_j \rho_{ij} C_{ij}(\theta)\}_{i,j=1}^{p}$, $\theta \in [0, \pi]$, where ρ_{ij} is a co-located correlation coefficient, and a completely monotone function $C_{ij}(\cdot)$ is of the form $C_{ij}(\theta) = \varphi_{[0,\pi]}(\theta; \boldsymbol{\lambda}_{ij}) = \int_{[0,\infty)} \exp(-\xi\theta)\mu(\mathrm{d}\xi; \boldsymbol{\lambda}_{ij})$, with μ being a positive and bounded measure and $\boldsymbol{\lambda}$ being a vector of parameters. Through the above mapping, they provided examples of cross-covariance functions. In particular, they showed the validity of the Matérn cross-covariances (Gneiting, Kleiber and Schlather, 2010, Apanasovich, Genton and Sun, 2012, Kleiber and Nychka, 2012) when used with the geodesic distance. However, we highlight that the Matérn covariance function with the geodesic distance is completely monotonic only for $\nu \in (0, 0.5]$ (Gneiting, 2013) and cross-covariances obtained through the above construction do not allow for differentiability at the origin. In practical applications, using the chordal distance may be preferable to using the geodesic distance for the flexible multivariate Matérn model due to this restriction on ν.

North, Wang and Genton (2011) presented derivations of isotropic covariance models across time on $\mathcal{S}^2$ for modeling global temperature fields. The forms were obtained from simple energy-balance climate models (i.e., white-noise-driven damped diffusion equations), which provided some physical insight into the dependencies. The resulting form on the plane belongs to the Matérn class with smoothness $\nu = 1$ (Whittle's covariance function) and that on $\mathcal{S}^2$ can be written as

$$C(\cos\theta) = \sigma^2 \sum_{k=0}^{\infty} \frac{2k+1}{4\pi} \frac{P_k(\cos\theta)}{\{\lambda^2 k(k+1)+1\}^2}, \tag{3}$$

where $\lambda > 0$ is the spatial range parameter and P_k is the Legendre polynomial of degree k.

Smoothing splines (Wahba, 1981) and spherical wavelets (Li, 1999, Oh and Li, 2004) have been explored to resolve the problem of modeling processes on $\mathcal{S}^2$ with multi-scale structures that can be spatially adaptive. These approaches are based on spherical harmonics. The spherical wavelets in Li (1999) and Oh and Li (2004) were constructed based on the Poisson kernel, which is analytic on $\mathcal{S}^2$.

3.3 Axially Symmetric Processes

Stationarity on $\mathcal{S}^2$ cannot be defined in the same way as in the planar domain. We can see from (1) that, even if a process is isotropic on $\mathcal{S}^2$ (i.e., isotropy with either the chordal or the geodesic distance is equivalent), its covariance function depends on the latitudes through L_1 and L_2, not the latitude lag between L_1 and L_2. Jones (1963) called a process on $\mathcal{S}^2$ *axially symmetric* if its covariance depends on the two latitude points and longitudinal lag. An isotropic process on $\mathcal{S}^2$ is axially symmetric, but axial symmetry does not imply stationarity in latitude in the usual sense.

Even if a spherical process is axially symmetric, it can exhibit various nonstationary covariance structures. First, as mentioned previously, the variance and the covariance may vary with the latitude in a complex manner. Second, it may be *longitudinally irreversible* (or *asymmetric*) for some L_1, L_2, ℓ_1, and ℓ_2:

$$\begin{aligned}&\operatorname{cov}\{Z(L_1, \ell_1), Z(L_2, \ell_2)\}\\ &\quad \neq \operatorname{cov}\{Z(L_1, \ell_2), Z(L_2, \ell_1)\}.\end{aligned}$$

In a similar manner, this asymmetry can be extended to the multivariate case and the space–time case. The covariance models proposed in Stein (2007), Jun and Stein (2008), and Jun (2011) are axially symmetric, but longitudinally irreversible and asymmetric (multivariate or space–time). However, any isotropic process on a sphere is axially symmetric, longitudinally reversible, and symmetric (multivariate or space–time). Recently, Ma (2016a, 2016b, 2016c, 2017) provided a general form of the covariance matrix structures for axially symmetric vector random fields, a series representation for a longitudinally reversible vector random field, and a time-varying vector random field.

4. CONSTRUCTIVE APPROACHES TO MODELING ON A SPHERE

4.1 Differential Operator

The key idea of the differential operator approach (Jun and Stein, 2007) was to apply a differential operator with respect to latitude and longitude to an isotropic process in the L^2 sense. If Z_0 is an isotropic process on $\mathcal{S}^2$, then

$$Z(L,\ell) = \left\{ A(L)\frac{\partial}{\partial L} + \frac{B(L)}{\cos L}\frac{\partial}{\partial \ell} \right\} Z_0(L,\ell). \tag{4}$$

Here, $A(L)$ and $B(L)$ are deterministic functions depending on latitudes, for example, they were modeled as finite linear combinations of spherical harmonics (Jun and Stein, 2008). In order for the form in (4) to be valid, the covariance function of Z_0 should be at least twice differentiable at the origin. As long as we have an explicit expression for the covariance function of Z_0, we can obtain an explicit expression for the covariance function of Z.

Jun (2011) extended the model in (4) to a multivariate setting for multivariate nonstationary and axially symmetric spatial processes on $\mathcal{S}^2$. Jun (2014) showed how the nonstationary Matérn model with spatially varying parameters (Kleiber and Nychka, 2012) can be coupled with the differential operator approach to achieve more flexible parametric cross-covariance models on $\mathcal{S}^2$. Jun (2014) also defined conditions on $A(L)$ and $B(L)$ so that the resulting process is mean-square continuous at the poles. Here, Matérn covariance functions are used with the chordal distance to model Z_0, but the differential operator approach can be applied to any process Z_0 regardless of the type of distance used to define the covariance function, as long as the model is valid and the derivatives of the covariance function exist at the origin.

4.2 Spherical Harmonic Representation

Stein (2007) used a series approach to model axially symmetric processes on $\mathcal{S}^2$. Let

$$Z(L,\ell) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} Y_{nm} \exp(im\ell) \tilde{P}_n^m(\sin L), \tag{5}$$

with Y_{nm} complex-valued random variables and the infinite sum converging in a mean-square sense. The polynomial P_n^m is the associated Legendre polynomial of degree n and order m, with $\tilde{P}_n^m$ being its normalized version, such that its squared integral on $[-1,1]$ is 1. Define $\delta_{ab} = 1$ if $a = b$, otherwise $\delta_{ab} = 0$. To achieve axial symmetry of Z in (5), we require that $\mathrm{E}(Y_{nm}) = \delta_{m0}\mu_n$, with $\sum_{n=0}^{\infty} \mu_n^2 < \infty$ and $\mathrm{cov}(Y_{nm}, Y_{n'm'}) = c_m(n,n')\delta_{mm'}$, under suitable restrictions on the complex-valued covariances $c_m(n,n')$. Further conditions on c_m are given in Stein (2007), who showed that the axially symmetric covariance function of the process in (5) is given by

$$\begin{aligned} &C(L,L',\ell) \\ &\quad = \sum_{m=-\infty}^{\infty} \sum_{n,n'=|m|}^{\infty} \exp(im\ell)\tilde{P}_n^m(\sin L) \\ &\qquad \cdot \tilde{P}_{n'}^m(\sin L') c_m(n,n'), \end{aligned} \tag{6}$$

where $c_{-m}(n,n') = c_m(n,n')^*$ and c_m^* is the complex conjugate of c_m. In practice, the infinite sums in (6) are truncated to a finite number of terms.

The spherical harmonic representation is a spectral representation with forms that allow conclusions to be drawn about local behaviors of the processes. To describe the large-scale features of the axially symmetric process it is useful to consider the covariance function using truncated series expansions. However, the computational burden associated with parameter estimation increases rapidly when accurately describing the small-scale behavior of a process.

4.3 Stochastic Partial Differential Equations

Using a stochastic partial differential equation (SPDE), Lindgren, Rue and Lindström (2011) showed the existence of an explicit link between Gaussian random fields and Gaussian Markov random fields. A random process on $\mathbb{R}^d$ with a Matérn covariance function is a solution to the SPDE:

$$(\kappa^2 - \Delta)^{\alpha/2} X(\mathbf{s}) = \phi \mathcal{W}(\mathbf{s}),$$

where $\mathcal{W}(\mathbf{s})$ is Gaussian white noise, Δ is the Laplace operator, and $\alpha = \nu + d/2$ (Whittle, 1963). Here, the

Matérn covariance function C has a shape parameter $\nu > 0$, a scale parameter $\kappa > 0$, and a variance parameter $\phi > 0$, with the parametrization $C(0) = \phi^2(4\pi)^{-d/2}\Gamma(\nu)\Gamma(\nu + d/2)^{-1}\kappa^{-2\nu}$.

Bolin and Lindgren (2011) extended the work of Lindgren, Rue and Lindström (2011) to construct a flexible class of models from a generalization of the SPDE models. This class maintains all of the desirable properties of the Markov approximated Matérn models, such as the computational efficiency, the convenient extensions to nonstationarity, and the applicability to data on general smooth manifolds. Nested SPDE models introduce nonstationarity by allowing the parameters κ_i, b_i, and $\mathbf{B}_i$ to vary spatially as

$$\left[\prod_{i=1}^{n_1}\{\kappa_i^2(\mathbf{s}) - \Delta\}^{\alpha_i/2}\right]X_0(\mathbf{s}) = \mathcal{W}(\mathbf{s}),$$

$$X(\mathbf{s}) = \left[\prod_{i=1}^{n_2}\{b_i(\mathbf{s}) + \mathbf{B}_i(\mathbf{s})^\top\nabla\}\right]X_0(\mathbf{s}),$$

where ∇ is the gradient, $b_i \in \mathbb{R}$, and $\mathbf{B}_i \in \mathbb{R}^d$. Here, $X(\mathbf{s})$ is the weighted sum of a Matérn field $X_0(\mathbf{s})$ and its directional derivative in the direction determined by the vector $\mathbf{B}_i$. This model is closely related to Jun and Stein (2008), but in this case the nested SPDEs inherently use the geodesic distance because they are directly defined on $\mathcal{S}^2$ (Bolin and Lindgren, 2011). This approach is computationally efficient in terms of CPU time and memory storage (Bradley, Cressie and Shi, 2015, 2016) via the Hilbert space approximation of the SPDE and the nested SPDE models, making it an attractive option for large data sets.

4.4 Kernel Convolution

Heaton et al. (2014) defined covariance functions on $\mathcal{S}^2$ in terms of the kernel convolution from Higdon (1998). They defined a spatial process as $W(\mathbf{s}) = \int_{\mathbf{u}\in\mathcal{S}^2} K\{\mathbf{u}|\boldsymbol{\eta}(\mathbf{s})\}Z(\mathbf{u})\,d\mathbf{u}$, where $K(\cdot)$ is a nonnegative, real-valued kernel for which $\int_{\mathbf{u}\in\mathcal{S}^2} K(\mathbf{u})\,d\mathbf{u} = 1$ and $Z(\mathbf{s})$ is a Gaussian white noise process with variance σ_W^2. The kernel satisfies symmetric conditions for all rotations, that is, $K(\mathbf{G}\mathbf{s}_0) = K(\mathbf{G}^{-1}\mathbf{s}_0)$, where $\mathbf{G} \in SO(3)$ is a real 3×3 orthogonal matrix from the special orthogonal group (Marinucci and Peccati, 2011). If K is stationary, then $K\{\mathbf{u}|\boldsymbol{\eta}(\mathbf{s}_2)\} = K\{\mathbf{G}^{-1}\mathbf{u}|\boldsymbol{\eta}(\mathbf{s}_1)\}$, and the covariance function is given by

$$\begin{aligned}&\operatorname{cov}\{W(\mathbf{s}_1), W(\mathbf{s}_2)\}\\ &\quad= \int_{\mathbf{u}\in\mathcal{S}^2} K\{\mathbf{u}|\boldsymbol{\eta}(\mathbf{s}_1)\}K\{\mathbf{u}|\boldsymbol{\eta}(\mathbf{s}_2)\}\,d\mathbf{u}\\ &\quad= \int_{\mathbf{u}\in\mathcal{S}^2} K\{\mathbf{u}|\boldsymbol{\eta}(\mathbf{s}_1)\}K\{\mathbf{G}^{-1}\mathbf{u}|\boldsymbol{\eta}(\mathbf{s}_1)\}\,d\mathbf{u},\end{aligned} \tag{7}$$

where $\boldsymbol{\eta}$ contains the common location-specific parameters (governing the orientation and dispersion) of the kernel. Heaton et al. (2014) considered the Kent distribution, which is a generalization of the von Mises–Fisher distribution defined on $\mathcal{S}^2$; see Kent (1982) for more details. The covariance function in (7) can be computed effectively by multiplying the Fourier coefficients using the convolution theorem (Heaton et al., 2014). By opting for rotation rather than geodesic distance, these models can capture geometrically anisotropic covariance structures.

Heaton et al. (2014) studied the relationship between the kernel and covariance functions through spherical harmonic decomposition. As in Section 4.2, their approximation via the discrete convolution method may not be appropriate for describing the small-scale properties of the process accurately. The kernel convolution via the Kent distribution could be useful for describing geometrically anisotropic covariance models and large-scale behaviors of the processes. However, in order to describe small-scale properties accurately, we need to try different kernels or approaches, for example, the von Mises–Fisher distribution leads to a smooth spatial field on $\mathcal{S}^2$ because its closed form of the covariance function has infinitely many derivatives.

4.5 Deformations

Sampson and Guttorp (1992) originally proposed the spatial deformation approach to modeling nonstationary spatial covariance structures on $\mathbb{R}^2$. The idea is that, given a nonstationary process on the plane, one can deform the geographical coordinate system so that the process is approximately isotropic in the deformed coordinate system. Das (2000) extended this method to $\mathcal{S}^2$ and proposed a new class of global anisotropic models. To generalize this method on $\mathcal{S}^2$, a method of deforming the global coordinate system and isotropic covariance functions are necessary. The covariance structure is expressed as a function of the geodesic distance between two locations on $\mathcal{S}^2$ in a bijective transformation of the coordinate system, $\operatorname{cov}\{Z(\mathbf{s}_1), Z(\mathbf{s}_2)\} = C[\theta\{g(\mathbf{s}_1), g(\mathbf{s}_2)\}]$, where C belongs to an isotropic class of parametric covariance functions on $\mathcal{S}^2$ and $g(\cdot)$ is a deformation function that expresses the spatial nonstationarity and anisotropy.

Here, the deformation should be restricted to the class of mappings from spheres onto spheres and the parallel latitude and longitude lines should not be allowed to cross. To specify functional forms of g, Das (2000) presented third-order spline functions, which

are quite flexible and capable of producing large deformations. Although this approach generates flexible models by composing deformation functions over and over, the computation is challenging for high-dimensional data. However, it is possible to draw conclusions about the deformed space because the deformation tends to compress (or stretch) the space among observation of high (or low) spatial correlations.

4.6 Multi-Step Spectrum

Castruccio and Stein (2013) and Castruccio and Genton (2014, 2016) introduced an alternative model of axial symmetry for regularly spaced data that results in a flexible covariance structure and computationally efficient model fitting. In their multi-step model, they separately considered the spectral process by latitudinal bands, fitted a covariance function across the longitudes, and then estimated the multi-band dependence. For a single latitude band at $L_m \in (-\pi/2, \pi/2)$, $m = 1, \ldots, M$, they proposed a parametric model for the spectrum $|f_{L_m}|^2$ at wave numbers $c = 0, \ldots, N-1$,

$$|f_{L_m}(c)|^2 = \phi_{L_m} / \{\alpha_{L_m}^2 + 4\sin^2(\pi c/N)\}^{\nu_{L_m}+1/2}, \tag{8}$$

which is similar to the Matérn spectrum with an overall variance ϕ_{L_m}, an inverse range parameter α_{L_m}, and a rate of decay of the spectrum ν_{L_m} at large wave numbers. Castruccio and Guinness (2017) relaxed the assumption of axial symmetry by allowing the properties of the process to vary over the earth's geography. They allowed the evolutionary transfer function $f_{L_m,\ell_n}(c)$ to depend on ℓ_n according to land and ocean covariates, such that $f_{L_m,\ell_n}(c) = f^1_{L_m}(c) b_{\text{land}}(L_m, \ell_n) + f^2_{L_m}(c)\{1 - b_{\text{land}}(L_m, \ell_n)\}$, where the component spectra has the parametric form (8) with different parameters for land and ocean, and b_{land} is an indicator function for different regimes. Recently, Jeong et al. (2017) extended this to also consider the topographical relief.

Multi-step models that are designed to evaluate the likelihood for parallel and distributed computing allow users to analyze very large data sets. For high-dimensional dependent data, high-performance computing can be integrated with the multi-step models using spatially varying parameters to analyze data of unprecedented size.

5. SIMULATED AND DATA EXAMPLES

5.1 Simulated Global Gaussian Processes with Different Covariance Structures

We perform simulations to show what processes with different covariance structures actually look like on unit spheres. We generate realizations from Gaussian processes with six different covariance structures. The first row of Figure 2 shows the realizations of isotropic processes with (a) the exponential covariance, (b) the Poisson kernel in (2), and (c) the covariance sum of the exponential covariance and Poisson kernel functions. We assume unit variances, but the spatial range (or shape) parameters, (β, r) are (0.157, 0.4), for (a), (b), and (c). For (c), we put more weight on the Poisson kernel with $\lambda = 0.7$. The second row of Figure 2 represents the realizations of nonstationary processes obtained from the different covariance structures. In (d), we assume the physics-based covariance in (3), applying different variance and range parameters over land and ocean, $(\sigma^2_{\text{land}}, \sigma^2_{\text{ocean}}, \rho_{\text{land,ocean}}) = (3, 1, 0.173)$ and $(\beta_{\text{land}}, \beta_{\text{ocean}}) = (0.235, 0.47)$. We set $\sigma^2(\ell) = 0.3 + 0.7P_3(\sin\ell)$ and $\beta(L) = 2 \times (90 - L)/180$. For both (e) and (f), we consider the exponential covariance models, but (e) has different variances based on the longitudinal bands and (f) has different ranges based on the latitudinal bands. For example, the effective range of spatial correlation is decreasing as the location changes from the South Pole to the North Pole in (f).

As the exponential covariance function has no derivative at the origin and the Poisson kernel is analytic, the realizations (a) and (b) show rough and smooth processes, respectively. The realization (c) obtained from the convex sum model shows a smoother realization than (a), but rougher than (b). By controlling the weight between the two models, we can control the smoothness of the process. For this reason, we apply the convex sum model to global temperature data in the following data example. The realization (d) shows that land areas are more variable than ocean areas. We clearly observe spatially varying structures depending on longitudinal and latitudinal bands in the realizations (e) and (f).

5.2 Global Surface Temperature Data

We illustrate covariance models using the geodesic distance on global surface temperature from the Geophysical Fluid Dynamics Laboratory (GFDL) model output. We use the monthly average surface temperature from December 2009 with 12,960 data locations across the earth. We filter out the mean structure of the data by spherical harmonic regression, with $\{Y_k^m(\sin L, \ell) | k = 0, 1, \ldots, K, m = -k, \ldots, k\}$ for $K = 12$, that is, $\mathbf{m}(L, \ell)^\top \boldsymbol{\beta}$ is $\sum_{k=0}^{12} \sum_{m=-k}^{k} f_k^m \cdot Y_k^m(\sin L, \ell)$. The estimated mean structure removes most of the large-scale spatial patterns, but some nonstationary land and ocean patterns remain (Figure 3).

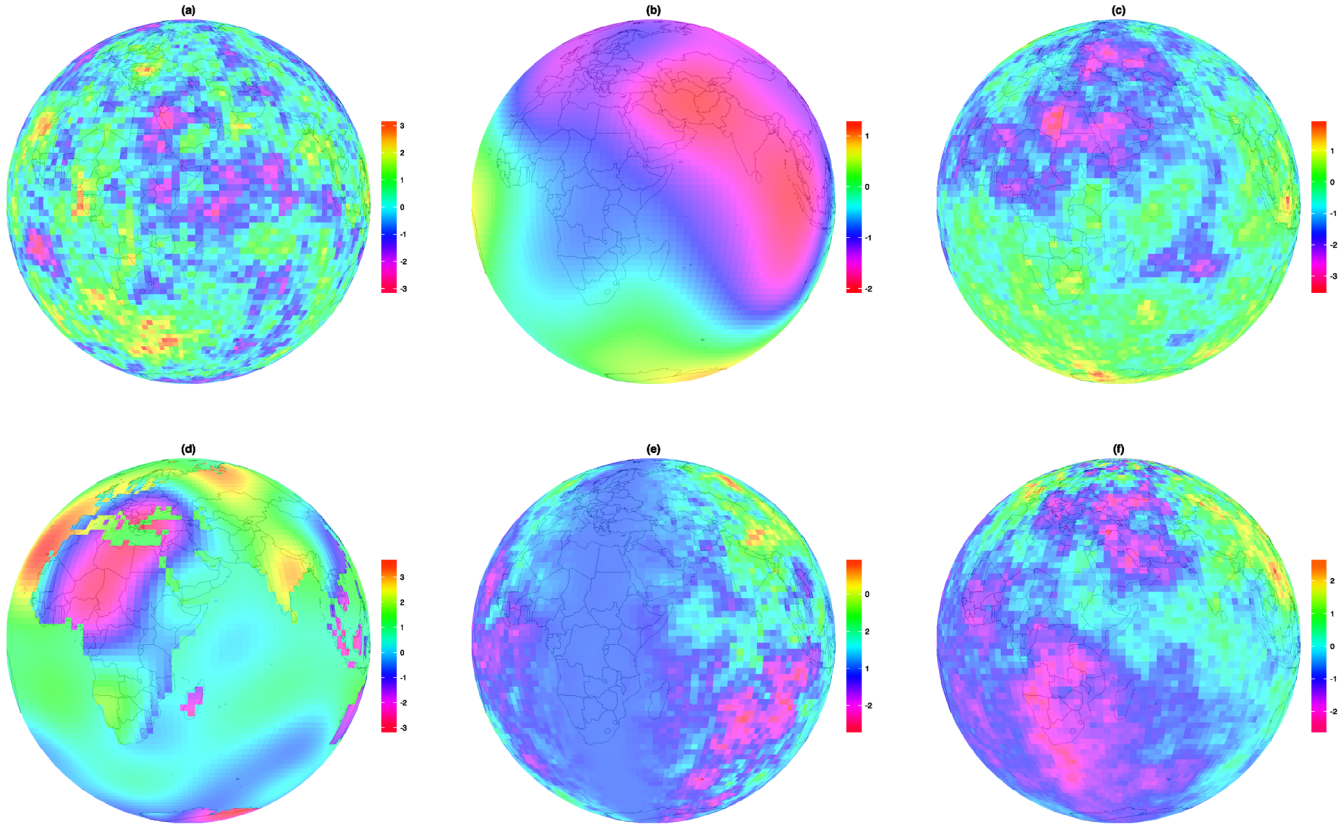

FIG. 2. *Realizations of Gaussian processes obtained from different covariance structures. Realizations of isotropic processes (first row) from* (a) *exponential covariance*, (b) *Poisson kernel, and* (c) *convex sum models. Realizations of nonstationary processes (second row) from* (d) *physics-based covariance with different parameters over land and ocean*, (e) *exponential covariance with different variances based on longitudinal bands, and* (f) *exponential covariance with different ranges based on latitudinal bands. The grid resolution for the realizations is* $2.5° \times 2°$.

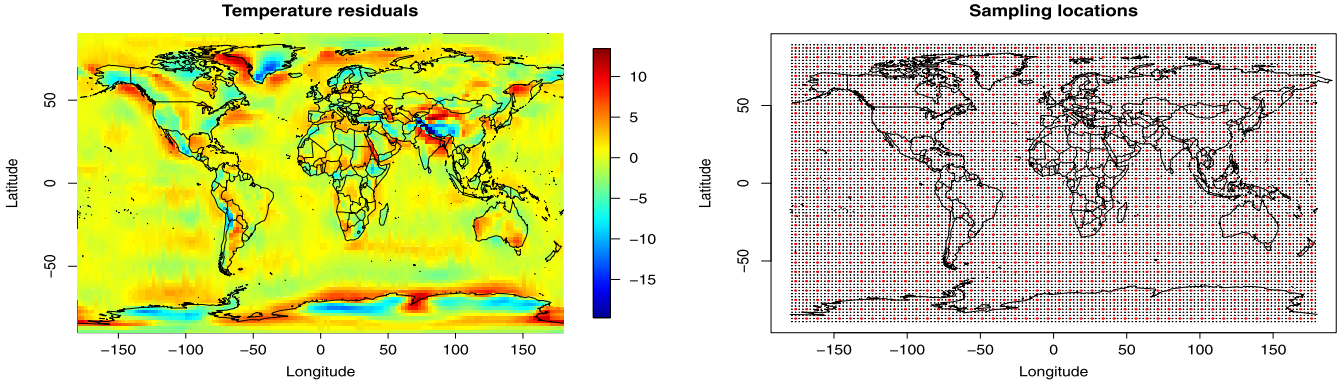

FIG. 3. *Residuals after removing the mean structure for average global surface temperature* (°*C*) *and sampling locations (red dots) for estimation and prediction. Equirectangular projection, in which the horizontal coordinate is the longitude and the vertical coordinate is the latitude, is used to construct plots.*

To model the nonstationary covariance structure, we consider the spatial deformation approach with isotropic covariance functions on $\mathcal{S}^2$ in terms of geodesic distance (Das, 2000), while incorporating the large-scale geographical indicators, land and ocean. For the isotropic covariance function, we use the convex sum model of the exponential and the Poisson kernel-like covariance functions, and multiply it by the Legendre polynomial of degree n to better account for some negative correlations for small distances and oscillating patterns around zero for large distances:

$$C(\theta) = \sigma^2 P_n(\cos\theta) \cdot \Bigg[\lambda \times \exp(-\theta/\beta) + (1-\lambda) \times \frac{(1-\alpha)^2}{\alpha(3-\alpha)}\left\{\frac{(1-\alpha)^2}{(1-2\alpha\cos\theta+\alpha^2)^{3/2}} - 1\right\}\Bigg],$$

where $n \in \mathbb{Z}$, $\lambda \in [0,1]$, $\sigma^2, \beta > 0$, and $\alpha \in (0,1)$. Das (2000) presented the basic 2D deformation functions of $g(L,\ell) = \{g_L(L,\ell), g_\ell(L,\ell)\}$ with third-order spline functions to find the new latitude and longitude:

$$g_L(L,\ell) = \begin{cases} \dfrac{1-b(\ell)}{90-\xi(\ell)}\{L-\xi(\ell)\}^2 + b(\ell)\{L-\xi(\ell)\} + \xi(\ell), \\ \qquad \xi(\ell) \le L \le 90, \\ -\dfrac{1-b(\ell)}{90+\xi(\ell)}\{L-\xi(\ell)\}^2 + b(\ell)\{L-\xi(\ell)\} + \xi(\ell), \\ \qquad -90 < L \le \xi(\ell), \end{cases} \tag{9}$$

where $\xi(\ell) = \alpha_\xi \exp\{-\cos^2(\ell-\eta_\xi)/\beta_\xi\}$ and $b(\ell) = \alpha_b \exp\{-\cos^2(\ell-\eta_b)/\beta_b\}$, and:

$$g_\ell(L,\ell) = \begin{cases} \dfrac{1-u(L)}{180-\eta(L)}\{\ell-\eta(L)\}^2 + u(L)\{\ell-\eta(L)\} + \eta(L), \\ \qquad \eta(L) \le \ell \le 180, \\ -\dfrac{1-u(L)}{180+\eta(L)}\{\ell-\eta(L)\}^2 + u(L)\{\ell-\eta(L)\} + \eta(L), \\ \qquad -180 < \ell \le \eta(L), \end{cases} \tag{10}$$

where $\eta(L) = \alpha_\eta \exp\{-\cos^2(L-\xi_\eta)/\beta_\eta\}$, and $u(L) = \alpha_u \exp\{-\cos^2(L-\xi_u)/\beta_u\}$. Here, $\alpha_b, \alpha_u \in [0,2]$, $\beta_\xi, \beta_b, \beta_\eta, \beta_u > 0$, $\alpha_\xi, \xi_\eta, \xi_u \in (-90, 90)$, $\alpha_\eta, \eta_\xi, \eta_b \in (-180, 180)$, $g_1(-90,\ell) = -90$, $g_1(90,\ell) = 90$, $g_2(L,180) = 180$, and $g_2(L,-180) = -180$. We focus on one step of the 2D deformation and include land and ocean covariates, so that $g^*(L,\ell) = \{g_L^*(L,\ell), g_\ell^*(L,\ell)\}$ can be expressed as $g_L^*(L,\ell) = g_L^1(L,\ell)1_{\text{land}}(L,\ell) + g_L^2(L,\ell)\{1 - 1_{\text{land}}(L,\ell)\}$ and $g_\ell^*(L,\ell) = g_\ell^1(L,\ell)1_{\text{land}}(L,\ell) + g_\ell^2(L,\ell)\{1 - 1_{\text{land}}(L,\ell)\}$, where $1_{\text{land}}(L,\ell)$ is an indicator function that modulates the contribution of land.

Four different covariances are calculated corresponding to (i) an isotropic model fit without deformations, (ii) a fit using a simple deformation with land and ocean covariates, where the deformations of one coordinate do not depend on the other coordinate [i.e., $\xi(\ell) = \xi$, $b(\ell) = b$, $\eta(L) = \eta$, and $u(L) = u$ in (9) and (10)], (iii) a fit using a flexible deformation with (9) and (10), where the deformations of a coordinate only depend on the other coordinate, and (iv) a fit using both the flexible deformation and geographical indicators, such as land and ocean. The parameters of the deformation and covariance functions are estimated using the maximum likelihood method.

To reduce the computational burden, we select 1440 regularly spaced locations (from 144×90 to 48×30), as shown in Figure 3. From these locations, we randomly select 90% to use as training data for the parameter estimation, and the remaining 10% (140 locations) are used as test data for the predictions. Conditional on the fitted covariance models, we compute the best linear unbiased predictions at each prediction location. Since each deformation model generates a different deformation space, we perform predictions on the deformed space and then compare the isotropic model with the deformation model in the prediction. We write Pred(A|B) to denote the prediction of case A on the deformed space of case B, and let Pred(B) = Pred(B|B), that is, Pred(i|ii) is the prediction from the isotropic model fit (i) on the deformed space generated by the deformation model fit (ii). As the criteria for comparison, we use the root mean squared error (RMSE), the mean absolute error (MAE), and the continuous ranked probability score (CRPS) to validate the prediction quality and average the predictions over their locations. We form 90% prediction intervals based on Gaussianity and report their empirical coverage for each model. We repeat this procedure 100 times for randomly chosen sets of estimation and prediction locations.

Table 1 contains the maximum log-likelihood values, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) values for the various models considered. All of the deformation models were superior to the isotropic model (i) in the

TABLE 1
Comparison of the log-likelihood values, AIC, and BIC averaged over 100 *replications for the convex sum models corresponding to different cases on the GFDL for surface temperature*

	(i)	(ii)	(iii)	(iv)
Number of parameters	5	13	17	29
Log-likelihood	−1488.0	−1425.9	−1410.5	**−1387.7**
AIC	2985.9	2877.8	2854.9	**2833.4**
BIC	3011.8	2945.0	**2942.8**	2983.3

model fit. The maximum log-likelihood value tends to increase as the number of parameters increases. In terms of model selection criteria, the flexible deformation model with land and ocean covariates (iv) is the best in AIC among the models under consideration. For BIC, similar values are observed between the deformation models (ii) and (iii). This suggests that, for geographical indicators such as land and ocean, these deformation models provide improved model fit.

Table 2 reports prediction results averaged over 100 replications for the isotropic model and the deformation models. Generally, we see better prediction scores using the deformation approach. Prediction scores of MAE and CRPS over the ocean are improved by 6–8% when using the simplest deformation model (ii) instead of the isotropic model (i). The more flexible deformation model (iii) shows similar patterns, and the most flexible model (iv) outperforms the isotropic model (i) in terms of prediction scores over both land and ocean. Both the widths of the prediction interval and the empirical coverage probabilities were smaller for the deformation models than for the isotropic model (i).

In general, we see that the deformation compresses the original space in regions of high spatial correlations, while stretching it in regions of relatively low spatial correlations. Figure 4 displays the fitted deformation coordinate systems. The deformation models with geographical indicators from cases (ii) and (iv) capture nonstationarity, based on whether the models are on land or ocean; however, they stretch land components in the Northern Hemisphere and contract the ocean components. The latitude −50 in the Southern Hemisphere is the most contracted because this area is comprised of mainly ocean with less variability than other areas. In addition, the deformation model (iii) identifies the high correlations at this latitude by contracting them.

Using simulations and data examples, we visualized realizations of the Gaussian processes obtained from different covariance structures on $\mathcal{S}^2$ and explored the fit of covariance models to climate data output. Because geophysical processes have complicated structures, the covariance models that we discussed might still be overly simplistic in many cases. Nevertheless, by investigating the structure of the data with covariance models ranging from basic to complex, a suitable and useful model can be found.

6. DISCUSSION

The models introduced in this paper deal with spherical data sets on a global scale and respect the curvature of the spherical domain with regard to the spatial dependence. However, the extension of covariance functions using the geodesic distance to the case of multivariate models, spatio-temporal models, nonstationary models, and non-Gaussian models is not widely understood. The development of various spatio-temporal covariance models based on the geodesic distance, as in Alegría et al. (2017b), is needed. For multivariate models, one natural direction is to explore physics-based cross-covariance models on $\mathcal{S}^2$, in which random fields

TABLE 2
Comparison of prediction scores, widths, and coverage probabilities of 90% *prediction intervals averaged over* 100 *replications on the GFDL for surface temperature*

	RMSE			MAE			CRPS				
Pred	**Land**	**Ocean**	**Total**	**Land**	**Ocean**	**Total**	**Land**	**Ocean**	**Total**	**Width**	**Coverage**
(i|ii)	**2.315**	1.271	**1.735**	1.494	0.788	1.042	**1.110**	0.710	0.855	6.11	93.1%
(ii)	2.382	**1.222**	1.746	**1.485**	**0.729**	**1.001**	1.117	**0.662**	**0.826**	5.55	91.6%
(i|iii)	**2.335**	1.272	**1.742**	1.529	0.804	1.065	**1.132**	0.692	0.851	5.81	92.5%
(iii)	2.374	**1.244**	1.750	**1.523**	**0.773**	**1.044**	1.133	**0.663**	**0.833**	5.46	91.6%
(i|iv)	2.371	1.360	1.803	1.551	0.840	1.096	1.155	0.729	0.882	5.83	91.5%
(iv)	**2.341**	**1.280**	**1.749**	**1.516**	**0.815**	**1.067**	**1.129**	**0.685**	**0.845**	5.45	90.7%

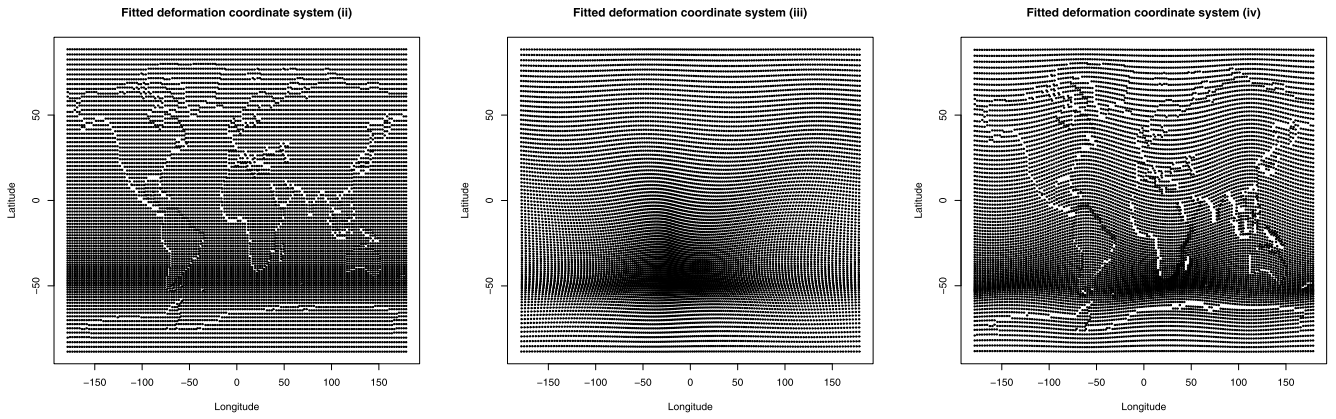

FIG. 4. *Maps of fitted deformations on the GFDL for surface temperature.*

generated from physical fields are constrained by physical laws. Cross-covariance models based on underlying physical processes are promising for applications in areas such as geophysical science, for example, modeling tangent vector fields on a sphere for zonal and meridional wind components (Fan et al., 2017). For flexible covariance modelings with real world applications, nonstationary models can be obtained by adapting spatially varying parameters in the covariance functions similarly to Jun (2014) and Poppick and Stein (2014). The Gaussian assumption makes a spatial model simple in structure and facilitates statistical predictions, but this assumption is often not supported by the data. To deal with this issue, we may consider non-Gaussian processes (Du et al., 2012, Ma, 2013a, 2013b, 2015, Du, Ma and Li, 2013), such as the skew-Gaussian (Alegría et al., 2017a) and the Tukey g-and-h random fields (Xu and Genton, 2017). These are promising for various applications, but their implementation remains unexplored. On the other hand, inference problems that are based on fixed-domain asymptotics are a challenge when modeling spatial processes. It is often infeasible to have consistent estimates from a single spatial realization (Zhang, 2004). Moreover, as outlined in Schaffrin (1993), consistent estimates of the covariance function on $\mathcal{S}^2$ exist only under the ergodic assumption, which means that all of the information characterizing a random field should be completely included in any of its realizations (Cressie and Wikle, 2011), with the random field Z being non-Gaussian. Few have explored the relationship between ergodicity and Gaussianity on $\mathcal{S}^2$ and more research is needed to allow better inference in spatial global models.

ACKNOWLEDGMENTS

This publication is based upon work supported by the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award No: OSR-2015-CRG4-2640.

REFERENCES

ALEGRÍA, A., CARO, S., BEVILACQUA, M., PORCU, E. and CLARKE, J. (2017a). Estimating covariance functions of multivariate Skew–Gaussian random fields on the sphere. *Spat. Stat.* To appear. Preprint. Available at arXiv:1612.03341.

ALEGRÍA, A., PORCU, E., FURRER, R. and MATEU, J. (2017b). Covariance functions for multivariate Gaussian fields evolving temporally over planet earth. Preprint. Available at arXiv:1701.06010.

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. *J. Amer. Statist. Assoc.* **107** 180–193. MR2949350

BANERJEE, S. (2005). On geodetic distance computations in spatial modeling. *Biometrics* **61** 617–625. MR2140936

BOLIN, D. and LINDGREN, F. (2011). Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. *Ann. Appl. Stat.* **5** 523–550. MR2810408

BRADLEY, J. R., CRESSIE, N. and SHI, T. (2015). Comparing and selecting spatial predictors using local criteria. *TEST* **24** 1–28. MR3314567

BRADLEY, J. R., CRESSIE, N. and SHI, T. (2016). A comparison of spatial predictors when datasets could be very large. *Stat. Surv.* **10** 100–131. MR3527662

CASTRUCCIO, S. and GENTON, M. G. (2014). Beyond axial symmetry: An improved class of models for global data. *Stat* **3** 48–55.

CASTRUCCIO, S. and GENTON, M. G. (2016). Compressing an ensemble with statistical models: An algorithm for global 3D spatio-temporal temperature. *Technometrics* **58** 319–328. MR3520661

CASTRUCCIO, S. and GUINNESS, J. (2017). An evolutionary spectrum approach to incorporate large-scale geographical descriptors on global processes. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **66** 329–344. MR3611690

CASTRUCCIO, S. and STEIN, M. L. (2013). Global space–time models for climate ensembles. *Ann. Appl. Stat.* **7** 1593–1611. MR3127960

CRESSIE, N. and JOHANNESSON, G. (2008). Fixed rank kriging for very large spatial data sets. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **70** 209–226. MR2412639

CRESSIE, N. and WIKLE, C. K. (2011). *Statistics for Spatio-Temporal Data*. Wiley, Hoboken, NJ. MR2848400

DAS, B. (2000). Global covariance modeling: A deformation approach to anisotropy. Ph.D. thesis, Univ. Washington, Seattle, WA. MR2701165

DU, J., MA, C. and LI, Y. (2013). Isotropic variogram matrix functions on spheres. *Math. Geosci.* **45** 341–357. MR3107159

DU, J., LEONENKO, N., MA, C. and SHU, H. (2012). Hyperbolic vector random fields with hyperbolic direct and cross covariance functions. *Stoch. Anal. Appl.* **30** 662–674. MR2946043

FAN, M., PAUL, D., LEE, T. C. M. and MATSUO, T. (2017). Modeling tangential vector fields on a sphere. *J. Amer. Statist. Assoc.* To appear. Preprint. Available at arXiv:1612.0806.

GNEITING, T. (1998). Simple tests for the validity of correlation function models on the circle. *Statist. Probab. Lett.* **39** 119–122. MR1652540

GNEITING, T. (1999a). Isotropic correlation functions on d-dimensional balls. *Adv. in Appl. Probab.* **31** 625–631. MR1742685

GNEITING, T. (1999b). Correlation functions for atmospheric data analysis. *Q. J. R. Meteorol. Soc.* **125** 2449–2464.

GNEITING, T. (2002). Nonseparable, stationary covariance functions for space–time data. *J. Amer. Statist. Assoc.* **97** 590–600. MR1941475

GNEITING, T. (2013). Strictly and non-strictly positive definite functions on spheres. *Bernoulli* **19** 1327–1349. MR3102554

GNEITING, T., KLEIBER, W. and SCHLATHER, M. (2010). Matérn cross-covariance functions for multivariate random fields. *J. Amer. Statist. Assoc.* **105** 1167–1177. MR2752612

GUINNESS, J. and FUENTES, M. (2016). Isotropic covariance functions on spheres: Some properties and modeling considerations. *J. Multivariate Anal.* **143** 143–152. MR3431424

HEATON, M. J., KATZFUSS, M., BERRETT, C. and NYCHKA, D. W. (2014). Constructing valid spatial processes on the sphere using kernel convolutions. *Environmetrics* **25** 2–15. MR3233740

HIGDON, D. (1998). A process-convolution approach to modelling temperatures in the North Atlantic Ocean. *Environ. Ecol. Stat.* **5** 173–190.

HITCZENKO, M. and STEIN, M. L. (2012). Some theory for anisotropic processes on the sphere. *Stat. Methodol.* **9** 211–227. MR2863609

HUANG, C., ZHANG, H. and ROBESON, S. M. (2011). On the validity of commonly used covariance and variogram functions on the sphere. *Math. Geosci.* **43** 721–733. MR2824128

JEONG, J. and JUN, M. (2015a). A class of Matérn-like covariance functions for smooth processes on a sphere. *Spat. Stat.* **11** 1–18. MR3311853

JEONG, J. and JUN, M. (2015b). Covariance models on the surface of a sphere: When does it matter? *Stat* **4** 167–182. MR3405399

JEONG, J., CASTRUCCIO, S., CRIPPA, P. and GENTON, M. G. (2017). Reducing storage of global wind ensembles with stochastic generators. Preprint. Available at arXiv:1702.01995.

JONES, R. H. (1963). Stochastic processes on a sphere. *Ann. Math. Stat.* **34** 213–218. MR0170378

JUN, M. (2011). Non-stationary cross-covariance models for multivariate processes on a globe. *Scand. J. Stat.* **38** 726–747. MR2859747

JUN, M. (2014). Matérn-based nonstationary cross-covariance models for global processes. *J. Multivariate Anal.* **128** 134–146. MR3199833

JUN, M. and STEIN, M. L. (2007). An approach to producing space–time covariance functions on spheres. *Technometrics* **49** 468–479. MR2394558

JUN, M. and STEIN, M. L. (2008). Nonstationary covariance models for global data. *Ann. Appl. Stat.* **2** 1271–1289. MR2655659

KENT, J. T. (1982). The Fisher–Bingham distribution on the sphere. *J. Roy. Statist. Soc. Ser. B* **44** 71–80. MR0655376

KLEIBER, W. and NYCHKA, D. (2012). Nonstationary modeling for multivariate spatial processes. *J. Multivariate Anal.* **112** 76–91. MR2957287

LANG, A. and SCHWAB, C. (2015). Isotropic Gaussian random fields on the sphere: Regularity, fast simulation and stochastic partial differential equations. *Ann. Appl. Probab.* **25** 3047–3094. MR3404631

LI, T.-H. (1999). Multiscale representation and analysis of spherical data by spherical wavelets. *SIAM J. Sci. Comput.* **21** 924–953. MR1755172

LINDGREN, F., RUE, H. and LINDSTRÖM, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 423–498. MR2853727

MA, C. (2012). Stationary and isotropic vector random fields on spheres. *Math. Geosci.* **44** 765–778. MR2956272

MA, C. (2013a). Mittag-Leffler vector random fields with Mittag–Leffler direct and cross covariance functions. *Ann. Inst. Statist. Math.* **65** 941–958. MR3105803

MA, C. (2013b). Student's t vector random fields with power-law and log-law decaying direct and cross covariances. *Stoch. Anal. Appl.* **31** 167–182. MR3007889

MA, C. (2015). Isotropic covariance matrix functions on all spheres. *Math. Geosci.* **47** 699–717. MR3368160

MA, C. (2016a). Isotropic covariance matrix polynomials on spheres. *Stoch. Anal. Appl.* **34** 679–706. MR3507186

MA, C. (2016b). Stochastic representations of isotropic vector random fields on spheres. *Stoch. Anal. Appl.* **34** 389–403. MR3488255

MA, C. (2016c). Time varying axially symmetric vector random fields on the sphere. *Random Oper. Stoch. Equ.* **24** 255–266. MR3574938

MA, C. (2017). Time varying isotropic vector random fields on spheres. *J. Theoret. Probab.* To appear.

MARINUCCI, D. and PECCATI, G. (2011). *Random Fields on the Sphere*: *Representation*, *Limit Theorems and Cosmological Applications*. *London Mathematical Society Lecture Note Series* **389**. Cambridge Univ. Press, Cambridge. MR2840154

NORTH, G. R., WANG, J. and GENTON, M. G. (2011). Correlation models for temperature fields. *J. Climate* **24** 5850–5862.

OH, H.-S. and LI, T.-H. (2004). Estimation of global temperature fields from scattered observations by a spherical-wavelet-based

spatially adaptive method. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **66** 221–238. MR2035768

POPPICK, A. and STEIN, M. L. (2014). Using covariates to model dependence in nonstationary, high-frequency meteorological processes. *Environmetrics* **25** 293–305. MR3258008

PORCU, E., BEVILACQUA, M. and GENTON, M. G. (2016). Spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere. *J. Amer. Statist. Assoc.* **111** 888–898. MR3538713

PORCU, E. and SCHILLING, R. L. (2011). From Schoenberg to Pick-Nevanlinna: Toward a complete picture of the variogram class. *Bernoulli* **17** 441–455. MR2797998

PORCU, E. and ZASTAVNYI, V. (2011). Characterization theorems for some classes of covariance functions associated to vector valued random fields. *J. Multivariate Anal.* **102** 1293–1301. MR2811618

SAMPSON, P. D. and GUTTORP, P. (1992). Nonparametric estimation of nonstationary spatial covariance structure. *J. Amer. Statist. Assoc.* **87** 108–119.

SCHAFFRIN, B. (1993). Biased kriging on the sphere? In *Geostatistics Tróia'* 92 (A. Soares, ed.) 121–131. Springer, New York.

SCHOENBERG, I. J. (1942). Positive definite functions on spheres. *Duke Math. J.* **9** 96–108. MR0005922

SCHREINER, M. (1997). On a new condition for strictly positive definite functions on spheres. *Proc. Amer. Math. Soc.* **125** 531–539. MR1353398

STEIN, M. L. (1999). *Interpolation of Spatial Data*: *Some Theory for Kriging*. Springer, New York. MR1697409

STEIN, M. L. (2007). Spatial variation of total column ozone on a global scale. *Ann. Appl. Stat.* **1** 191–210. MR2393847

WAHBA, G. (1981). Spline interpolation and smoothing on the sphere. *SIAM J. Sci. Statist. Comput.* **2** 5–16. MR0618629

WHITTLE, P. (1963). Stochastic processes in several dimensions. *Bull. Int. Stat. Inst.* **40** 974–994. MR0173287

XU, G. and GENTON, M. G. (2017). Tukey g-and-h random fields. *J. Amer. Statist. Assoc.* **112** 1236–1249.

YADRENKO, M. Ĭ. (1983). *Spectral Theory of Random Fields*. Optimization Software, Inc., Publications Division, New York. Translated from the Russian. MR0697386

YAGLOM, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions. Vol. I*: *Basic Results*. Springer, New York. MR0893393

ZHANG, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *J. Amer. Statist. Assoc.* **99** 250–261. MR2054303