A class of Matérn-like covariance functions for smooth processes on a sphere

Jaehong Jeong, Mikyoung Jun

Department of Statistics, Texas A&M University, College Station, TX, United States

ARTICLE INFO

Article history:
Received 21 March 2014
Accepted 3 November 2014
Available online 13 November 2014

Keywords:
Climate model outputs
Covariance function
Matérn model
Process on a sphere
Smoothness of a process

ABSTRACT

There have been noticeable advancements in developing parametric covariance models for spatial and spatio-temporal data with various applications to environmental problems. However, literature on covariance models for processes defined on the surface of a sphere with great circle distance as a distance metric is still sparse, due to its mathematical difficulties. It is known that the popular Matérn covariance function, with smoothness parameter greater than 0.5, is not valid for processes on the surface of a sphere with great circle distance. We introduce an approach to produce Matérn-like covariance functions for smooth processes on the surface of a sphere that are valid with great circle distance. The resulting model is isotropic and positive definite on the surface of a sphere with great circle distance, with a natural extension for nonstationarity case. We present extensive numerical comparisons of our model, with a Matérn covariance model using great circle distance as well as chordal distance. We apply our new covariance model class to sea level pressure data, known to be smooth compared to other climate variables, from the CMIP5 climate model outputs.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

In geophysical and environmental sciences, datasets often come in a global scale. Covariance models for global data need to be positive definite on the surface of a sphere and their natural distance metric is great circle distance. However, literature on covariance modeling on the surface of
a sphere with great circle distance is scarce, since it is mathematically challenging. Schoenberg (1942) presented a classical representation for positive definite functions with great circle distance in terms of Gegenbauer polynomials. Yadrenko and Balakrivan’ (1983) and Yaglom (1987) presented a simple but useful idea of restricting valid isotropic covariance functions in \( \mathbb{R}^2 \) on the surface of a sphere, which requires the use of chordal distance rather than great circle distance. Huang et al. (2011) examined the validity of several parametric isotropic covariance functions with great circle distance on the surface of a sphere. Some parametric covariance functions on the surface of a sphere are derived from various constructional approaches (Jun and Stein, 2007, 2008; Huang et al., 2011; Du and Ma, 2012; Du et al., 2013; Gneiting, 2013).

The Matérn covariance function is given as

\[
K(t; \sigma^2, \alpha, \nu) = \frac{\pi^{1/2} \sigma^2 (\alpha^{2\nu})}{2^{\nu-1} \Gamma(\nu + 1/2)} (t/\alpha)^\nu \mathcal{K}_\nu(t/\alpha), \quad t \geq 0,
\]

where \( \sigma^2 > 0, \alpha > 0 \), and \( \nu > 0 \) represent marginal variance, spatial range, and smoothness parameters, respectively. Here, \( \mathcal{K}_\nu \) is the modified Bessel function of the second kind of order \( \nu \). The larger \( \nu \) is, the smoother the resulting Matérn field is, and smoothness means mean square smoothness (Stein, 1999). In particular, the Matérn field is at least \( m \) times mean square differentiable if and only if \( \nu > m \), due to the fact that

\[
K(t) = \begin{cases} 
\sum_{j=0}^{m} a_j t^{2j} - \frac{\pi \sigma^2}{\Gamma(2\nu + 1)} t^{2\nu} \log(t) + O(t^{2m+2}), & \text{as} \ t \downarrow 0, \quad \text{if} \ m < \nu < m + 1, \\
\sum_{j=0}^{m} b_j t^{2j} + \frac{2(-1)^m \sigma^2}{(2m + 2)!} t^{2m+2} \log(t) + O(t^{2m+2}), & \text{as} \ t \downarrow 0, \quad \text{if} \ \nu = m + 1,
\end{cases}
\]

for appropriate constants \( a_j \) and \( b_j \), \( j = 1, \ldots, m \), depending on \( \sigma^2, \nu \), and \( \alpha \) (Stein, 1999). The Matérn class is positive definite in \( \mathbb{R}^d \) for any \( d \in \mathbb{N} \) with the Euclidean distance, i.e., chordal distance on the surface of a sphere (Stein, 1999).

According to Miller and Samko (2001) and Gneiting (2013), the Matérn class using great circle distance is valid if and only if the smoothness parameter satisfies \( 0 < \nu \leq 0.5 \) for 1, 2, and 3 dimensional spheres, and thus it is not mean square differentiable at the origin. This implies that the Matérn class with great circle distance cannot be adapted for smooth processes on the surface of a sphere.

We propose a new class of Matérn-like covariance functions with great circle distance on the surface of a sphere. Resulting covariance function guarantees positive definiteness and is designed to model smooth processes on the surface of a sphere. The new approach integrates a non-differentiable process over a small neighborhood on the surface of a sphere to make it differentiable, and uses finite summation approximation along with practical algorithm for the approximation. Furthermore, by considering local integration vary over space, it can be easily adapted for nonstationary processes.

Hansen et al. (2011) provided a flexible framework with Lévy particles for modeling and simulating star-shaped random particles in three dimension. Our covariance functions may be connected to this framework with a certain choice of kernel functions. However, parameter estimation is not discussed in details in the paper and it is not clear how practical the proposed approach is.

Recently, Guinness and Fuentes (2013) proposed the, so called, circular Matérn covariance function in terms of a Fourier series, which is valid on the surface of a sphere with great circle distance. They argued that the circular Matérn covariance function is \( m \) times mean square differentiable by defining the notion of mean square differentiable processes on the surface of a sphere. It has a closed-form only when the smoothness parameter is a half-integer. For other smoothness values, it is expressed as an infinite summation. Therefore, in practical application, we may need to truncate the summation at finite terms. To deal with truncation issue for arbitrary smoothness values, they addressed the approximation theorems that provide guarantees about positive definiteness. They also give computationally efficient methods.

The rest of the paper is organized as follows. Section 2 presents the new method proposed. It provides an illustration of the method on \( S^1 \) and provides details on implementation on \( S^d \) for
\(d = 1, 2, 3\). Then we present four simulation studies on \(S^2\) in Section 3. Section 4 gives real application results to a sea level pressure data. Finally, Section 5 provides some discussion and future work.

2. Method

Our goal is to construct a class of Matérn-like covariance functions for smooth processes on the surface of a sphere, which is valid with great circle distance. To avoid the hassle of verifying the positive definiteness of proposed covariance models, we start from a homogeneous process whose covariance structure is given by a Matérn class with \(0 < \nu \leq 0.5\). Our key idea is to integrate such process over a small neighborhood on the surface of a sphere, which results in a smoother process.

Suppose \(Z(\mathbf{s}), \mathbf{s} \in S^d, d = 1, 2, 3\) (\(S^d\) is the surface of a sphere with radius \(R\) in \(\mathbb{R}^{d+1}\)), is a homogeneous process on the surface of a sphere, \(\text{cov}(Z(\mathbf{s}), Z(\mathbf{t})) = \mathbf{K}(\theta; \sigma^2, \alpha, \nu)\) with \(\theta = \theta(\mathbf{s}, \mathbf{t}) = \arccos((\mathbf{s}, \mathbf{t}))\), and \(K\) given in (1). Then let us define a new process on \(S^d\),

\[
W_{\epsilon}(\mathbf{s}) = \frac{1}{\Delta_{\epsilon}} \int_{\delta_{\epsilon}(\mathbf{s})} Z(\mathbf{t}) \, d\mathbf{t},
\]

where \(\delta_{\epsilon}(\mathbf{s}) = \{\mathbf{u} \in S^d : \theta(\mathbf{s}, \mathbf{u}) \leq \epsilon\}\) for some \(\epsilon > 0\) and the integral is done on the surface of a sphere. Further, \(\Delta_{\epsilon}\) denotes the area of \(\delta_{\epsilon}(\cdot)\) on the surface of the sphere (with radius, \(R\)). Then, the covariance of \(W_{\epsilon}\) is given by,

\[
\text{cov}(W_{\epsilon}(\mathbf{s}), W_{\epsilon}(\mathbf{t})) = \frac{1}{\Delta_{\epsilon}^2} \int_{\delta_{\epsilon}(\mathbf{s})} \int_{\delta_{\epsilon}(\mathbf{t})} \mathbf{K}(\theta(\mathbf{u}, \mathbf{v}); \sigma^2, \alpha, \nu) \, d\mathbf{u} \, d\mathbf{v},
\]

where \(\sigma^2 > 0, \alpha > 0, 0 < \nu \leq 0.5\), and \(\epsilon > 0\) represent marginal variance, spatial range, smoothness, and integration range parameters for \(Z\), respectively. For simplicity, we assume that mean of \(Z\) is zero, and thus mean of \(W_{\epsilon}\) is zero. If we assume that \(Z\) follows Gaussian distribution (which is often done for spatial and spatio-temporal data in practice), then the distribution of \(W_{\epsilon}\) is also Gaussian. This is due to the fact that all linear functionals of a Gaussian process, such as derivatives or integrals of Gaussian processes, are Gaussian processes (Parzen, 1999). Note that we normalize the process in (3) by \(\Delta_{\epsilon}\), so that variance of \(W_{\epsilon}\) is the same as that of \(Z\). From now on, we drop \(\epsilon\) in \(W_{\epsilon}\).

The model proposed in (3) may seem similar to a kernel averaged process with a top-hat filter, a function shaped similar to a density function of continuous uniform distribution. Along that line, one may consider \(\epsilon\) as a bandwidth of such kernel. Note that Heaton et al. (2014) used a specific kernel to model processes on the surface of a sphere. However, the motivation of integrating a process in (3) is quite different from kernel averaging. The main purpose for the integration in (3) is to achieve a smoother process. One may fix \(\epsilon\) at a very small value to achieve a smooth \(W\). In fact, moderate or large values of \(\epsilon\) may result in undesirably strong spatial dependence of \(W\) at a small scale, which may cause numerical problems in dealing with covariance matrix of \(W\). Adding a kernel function in (3) results in a complex form of covariance function in (4), which is not desirable in our case. To achieve an even smoother process, we may simply keep repeating the integrals in (3). See (10) in Section 3.2 for further details.

2.1. Illustration on \(S^1\)

In \(S^1\), the analytic calculation of the double integrals in (4) is relatively simple. Consider a process \(W(s) = \int_{-\epsilon}^{\epsilon} Z(u) \, du/(2\epsilon)\) on \(S^1\). In particular, when \(K\) is an exponential function with range \(\alpha\), we can easily obtain an explicit expression for the covariance function of \(W\). Some tedious calculation gives us \(\langle \theta = \theta(s, t) = |t - s|\rangle:\n
(2\epsilon)^2 \text{cov}(W(s), W(t)) = \begin{cases} 
\alpha(4\epsilon - 2\theta) + \alpha^2 \left( -2e^{-\theta/\alpha} + e^{\theta/\alpha} + e^{-\theta - 2\alpha} \right), & \text{if } \theta \leq 2\epsilon, \\
\alpha^2 e^{-\theta/\alpha} (e^{2\alpha/\epsilon} + e^{-2\alpha/\epsilon} - 2), & \text{if } 2\epsilon < \theta < \pi - \epsilon, \\
\alpha^2 \left( -2e^{-\theta/\alpha} + 2e^{-\theta - \pi} + \left( e^{\theta/\alpha} + e^{-\theta/\alpha} \right) (e^{2\beta/\alpha} - 1) \right), & \text{if } \pi - \epsilon \leq \theta \leq \pi.
\end{cases}
\]
Thus, when the estimate of $\nu$ is not larger than the truth, we tend to get smaller estimates of $\alpha$ and $\epsilon$. This result is not surprising, since MG is not suitable for smooth processes due to the distortion at large distances. We found that there is negative association between $\epsilon$ and $\nu$ in (5), and compared fitted results using the true covariance model (we call it MI$^*$) to the original correlation function of $Z$, $\rho(\theta) = e^{-\theta/\alpha}$, and the integrated correlation functions of $W$ for $\epsilon = 0.05$, $0.15$, and 0.3. Variabilities of the integrated correlation functions decrease slower near the origin than the original correlation function, and the integrated correlation functions look smoother at the origin compared to the original correlation function of $Z$.

As a sanity check, we simulated Gaussian random fields on $S^1$ using the proposed covariance model in (5), and compared fitted results using the true covariance model (we call it MI$^*$). We use $\ast$, since by using exponential model, $\nu$ for $Z$ is fixed at 0.5. See Section 3.1 for details on the notation as well as Matérn covariance functions with chordal distance (MC) and great circle distance (MG). We considered the spatial range $\alpha = \pi/2$ and $\pi$ and $\epsilon = 0.03$. We randomly selected 100 locations on a unit circle, and among them, we used 90 randomly selected data points for parameter estimation and the remaining 10 data points for prediction. We repeated this procedure 100 times.

Table 1 shows sample means and standard deviations for the model parameter estimates through maximum likelihood estimation, RMSPE, MAE, and CRPS under each model. We observed that the range parameter $\alpha$ is underestimated for MC and overestimated for MG. While all models have comparable performance in prediction, the true model, MI$^*$, outperforms MG, in terms of maximum log-likelihood values. This result is not surprising, since MG is not suitable for smooth processes due to the restriction on smoothness parameter ($\nu \leq 0.5$). On the other hand, MC performs reasonably well despite the distortion at large distances. We found that there is negative association between the estimates of $\epsilon$ and $\alpha$. This is due to the fact that larger $\epsilon$ values tend to result in larger effective range. Thus, when the estimate of $\epsilon$ is larger than the truth, we tend to get smaller estimates of $\alpha$ to compensate for it. We tested cases with smaller true $\alpha$ values and there was no noticeable change, in terms of how the three models compare.

When the smoothness of $Z$ is $0 < \nu \leq 0.5$, due to (2), we achieve the following result. As $\theta = |t - s| \downarrow 0$,

$$\text{cov}\{W(s), W(t)\} = a_0 + a_1 \theta^2 + c_0 \theta^{2(\nu+1)} + O(\theta^4),$$

for $c_0 = \frac{-\pi \sigma^2}{2 \pi^2 F[2(\nu+1)+1] \sin[(\nu+1)\pi]}$, and appropriate constants $a_0$, $a_1$, depending on $\sigma^2$, $\nu$, $\alpha$, and $\epsilon$. Theorem 2 of Stein (1999) and (7) imply that the covariance function of $W$ has 2 derivatives at the origin.

2.2. Practical approximation of integral for $S^d$

In general, analytically evaluating the integral in (4) on the surface of the sphere is tricky, except for those simple situations presented in Section 2.1. Instead, we use a finite approximation of the
The exponential correlation function with marginal variance 1 and spatial range parameter $\alpha = \pi$ (black line), integrated exponential correlation functions given in (5) with $\epsilon = 0.05, 0.15$ and $0.3$ (straight line) and its approximation as in (8) (dotted line).

integrals as in the following $\mathbf{s}, \mathbf{t} \in \delta^d, \ d = 1, 2, 3$:

$$\text{cov}\{W(\mathbf{s}), W(\mathbf{t})\} \approx \frac{1}{N^2} \sum_{\mathbf{u}_i \in \tilde{\delta}_\epsilon(\mathbf{s})} \sum_{\mathbf{v}_j \in \tilde{\delta}_\epsilon(\mathbf{t})} K(\theta(\mathbf{u}_i, \mathbf{v}_j)), \tag{8}$$

where $K(\cdot)$ is a Matérn covariance function in (1) with $0 < \nu \leq 0.5$ and $i, j = 1, \ldots, N$. Here, $\tilde{\delta}_\epsilon(\mathbf{s})$ is a set of a finite number of locations defined as in the following. First, consider $n_c$ number of circles with a center $\mathbf{s}$ and radii, $\epsilon, \epsilon/2, \epsilon/4, \ldots, \epsilon/2^{n_c-1}$. Then on each circle, we consider $n_p$ as many equally spaced (in terms of angles) points. Thus points on each circle are $2\pi/n_p$ radians apart. Then the set $\tilde{\delta}_\epsilon(\mathbf{s})$ consists of $\mathbf{s}$ and $n_p$ points on the circles; $\tilde{\delta}_\epsilon(\mathbf{s})$ contains $n_c \times n_p + 1$ points. Fig. 2 gives an example of the choice of the finite number of locations used in the approximation in (8), and there are 8 points on each circle. Algorithm 1 returns a point that is on a circle with specified radius (from the center) and on a circle, at a specified angle apart (clockwise) from the vertical line. When we compute the covariance between two locations on the surface of a sphere, we use a finite approximation of the integrals in (8) by averaging covariances among all possible pairs of $n_c \times n_p + 1$ locations. Note that for any $n_c$ and $n_p$, (8) gives a positive definite covariance function, as it is simply a summation of a finite number of positive definite functions.

Instead of what is proposed in Algorithm 1, one may choose $\mathbf{u}_i$’s and $\mathbf{v}_j$’s randomly on each circle. However, for small $n_c$ and $n_p$, randomly selecting locations of $\mathbf{u}_i$’s and $\mathbf{v}_j$’s may not be effective. When $n_p$ is small, most points may end up close to each other, and thus the finite summation in (8) may approximate only a small part of the domain of the integral. Our code to calculate the approximation in (8), originally written in C, is available for usage in R, and is available at http://www.stat.tamu.edu/~mjunkjeong_jun.html.

2.3. Choice of $\epsilon$, $n_c$, and $n_p$

The proposed approach requires choices of $\epsilon$, the number of circles $n_c$, and the number of points on each circle $n_p$. When $\epsilon = 0$, the integrated covariance function reduces to the covariance function of $Z$, which is suitable for rough processes. To choose $\epsilon$ adaptively, we estimate $\epsilon$ along with the parameters of the Matérn covariance function, through the maximum likelihood method. As shown in the example
The locations used for the approximation, for the centers (−35°, 40°), (70°, −10°), and (−20°, −40°) with great circle distance 2000 (km) as a radius, obtained from Algorithm 1. The number of circles and the number of points on each circle are 3 and 8, respectively.

Algorithm 1 Great circle waypoint. This algorithm returns a point that is a specified great circle distance apart from the center and at a specified angle apart from the center (clockwise) ($l$: longitude, $L$: latitude, $\theta$: great circle distance, and $\phi$: angle).

\begin{verbatim}
procedure NEW_LOC($l_1, L_1, \theta, \phi$)
    convert degrees to radians
    $l_1 = l_1 \times \pi / 180$, $L_1 = L_1 \times \pi / 180$, $\phi = \phi \times \pi / 180$
    convert great circle distance to radian
    $d = \theta / R$ ( $R$ : radius of the sphere)
    $a = \sin(d) \times \sin(\phi)$
    $b = \cos(L_1) \times \cos(d) - \sin(L_1) \times \sin(d) \times \cos(\phi)$
    if ($b = 0$) then
        $l_2 = l_1$
    else
        $l_2 = l_1 + \arctan2(a, b)$
    end if
    convert radians to degrees
    $l_2 = l_2 \times 180 / \pi$, $L_2 = L_2 \times 180 / \pi$
    return $l_2$ and $L_2$
end procedure
\end{verbatim}

in Section 2.1, the estimation of $\epsilon$ along with other covariance parameters works well with small $n_c$ and $n_p$.

The choice of $n_c$ and $n_p$ involves a trade-off between the computational cost and accuracy of approximation. According to our experience, small numbers, such as $n_c = 1$ or 2 and $n_p = 8$ or so, give a reasonable approximation in terms of model fit and spatial prediction. For example, we tested the approximation for the model in Section 2.1. For each location $s \in S^1$, we used 6 points, $s \pm \epsilon$, $s \pm \epsilon / 2$, and $s \pm \epsilon / 4$, to approximate $\text{cov}(W(s), W(t))$. The dotted lines of Fig. 1 represent approximated correlation functions with $\epsilon = 0.05$, 0.15, and 0.3. True lines from (5) and their approximation through finite summations are nearly identical.

2.4. Extension for nonstationarity

In real applications, the assumption of isotropy may be too limited. Data in geophysical applications often exhibit strong dependence of covariance structure on latitude (Jun and Stein, 2008), and the spatial range or smoothness parameters may vary across space (Kleiber and Nychka, 2012; Jun, 2014). A straightforward extension of our model for nonstationarity may be letting $\epsilon$ vary over space. For example, we may allow $\epsilon$ to be different over the land and the sea for some climate variables, or to
3. Simulation studies

We present four simulation studies for the case on $\delta^2$, and compare the proposed covariance model to Matérn covariance models with chordal distance as well as great circle distance. We also present comparison among the models using great circle distance in terms of differentiability at the origin. In the first example (Section 3.1), the truth is generated from a Matérn model with chordal distance. In the second example (Section 3.2), the truth is generated from compactly supported covariance functions with great circle distance. In the third example (Section 3.3), a nonstationary Matérn covariance model using chordal distance is used as the truth. In the final example (Section 3.4), non-differentiable model, as well as differentiable models (once and twice differentiable) using great circle distance are used as the truth.

3.1. Simulation I

We consider a Gaussian random field on the surface of the Earth (with radius $R = 6371$ (km)) with mean 0 and a Matérn covariance function in (1) with chordal distance. Note that chordal distance between the two locations, $(L_1, l_1)$ and $(L_2, l_2)$, on $S^2$ ($L$ and $l$ denote latitude and longitude, respectively) is given by

$$\text{ch}(L_1, L_2, l_1 - l_2) = 2R \left\{ \sin^2 \left( \frac{L_1 - L_2}{2} \right) + \cos L_1 \cos L_2 \sin^2 \left( \frac{l_1 - l_2}{2} \right) \right\}^{1/2}.$$ 

Great circle distance between the two locations then is given by $\theta = \text{gc}(L_1, L_2, l_1 - l_2) = 2R \cdot \arcsin \left\{ \text{ch}(L_1, L_2, l_1 - l_2)/(2R) \right\}$. We set $\sigma^2 = 1$, $\alpha = 1000$ (km), and $\nu = 0.5$, 0.75, 1.5, 2.5. We randomly selected 500 locations over the surface of the Earth. Among 500 data points, we used 450 randomly chosen data points as training data for parameter estimation, and remaining 50 data points as test data for prediction. If we randomly select locations on the Earth based on angles, we may end up with clusters near the poles, since two points with the same angular distance apart near the poles are much closer to each other, compared to two points near the equator. To obtain spatially more regular sample locations over the globe, we sampled locations with percentages depending on the latitude, roughly proportional to the area of each latitude band. That is, we use the percentages, (25.88, 24.12, 20.71, 15.89, 9.99, 3.41), for latitudes, ($\pm[0^\circ, 15^\circ]$, $\pm(15^\circ, 30^\circ]$, $\pm(30^\circ, 45^\circ]$, $\pm(45^\circ, 60^\circ]$, $\pm(60^\circ, 75^\circ]$, $\pm(75^\circ, 90^\circ]$), respectively.

We used the maximum likelihood estimation method for each model with the optim function in R. We compared fitted results from Matérn model with chordal distance (MC) to those from Matérn model with great circle distance (MG), as well as those from the proposed model (MI). For MI, we use the notation MIx,y to note that $n_c = x$ and $n_p = y$ used in the approximation of the integrals. Often for the proposed model MI, the estimated $\nu$ values are close to 0.5 (particularly for smooth processes). Thus, we fix $\nu = 0.5$ to save computational time. For such cases, we denote by MIx,$y^\nu$. For MC, $\nu > 0$ is enough to guarantee positive definiteness, and for MG and MI, we need $0 < \nu \leq 0.5$. We compared the models in terms of prediction, and used RMSPE and two popular scoring rules (Gneiting and Raftery, 2007), the mean absolute error (MAE) and the continuous ranked probability score (CRPS), as the criteria for comparison. We repeat each simulation 100 times.

Table 2 displays sample means and standard deviations for the model parameter estimates under each model. When $\nu = 0.5$, the estimates of range parameter $\alpha$, and the smoothness parameter $\nu$ are close to the true Matérn covariance model using chordal distance. Maximum loglikelihood values and prediction measures for each model, are similar, as expected. When $\nu \geq 1.5$, we observe that the range parameter $\alpha$ is overestimated and the smoothness parameter $\nu$ is estimated close to 0.5 for MG. From Table 2, we observe that MG has the largest RMSPE, MAE, and CRPS values for $\nu \geq 1.5$, and
Table 2
(Simulation Example I) Sample means and standard deviations of parameter estimates, maximum loglikelihood values, and prediction measures for each model (100 cases). True covariance function is the Matérn model with chordal distance. We set $\sigma^2 = 1$, $\alpha = 1000$, and $\nu = 0.5, 0.75, 1.5, 2.5$.

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>Model</th>
<th>$\hat{\epsilon}$</th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\nu}$</th>
<th>Max. loglik</th>
<th>RMSPE</th>
<th>MAE</th>
<th>CRPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>MC</td>
<td>1035.220(256.586)</td>
<td>0.512(0.088)</td>
<td>-121.483(14.790)</td>
<td>0.714(0.080)</td>
<td>0.562(0.064)</td>
<td>0.398(0.042)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG</td>
<td>1106.760(210.192)</td>
<td>0.471(0.046)</td>
<td>-121.697(14.807)</td>
<td>0.713(0.081)</td>
<td>0.562(0.064)</td>
<td>0.398(0.043)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MI1.8</td>
<td>19.827(28.344)</td>
<td>0.445(0.085)</td>
<td>-121.338(14.739)</td>
<td>0.714(0.081)</td>
<td>0.563(0.065)</td>
<td>0.399(0.043)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MI2.8</td>
<td>23.285(33.759)</td>
<td>0.445(0.093)</td>
<td>-121.309(14.719)</td>
<td>0.714(0.080)</td>
<td>0.563(0.064)</td>
<td>0.399(0.043)</td>
<td></td>
</tr>
<tr>
<td>0.75</td>
<td>MC</td>
<td>1015.962(194.789)</td>
<td>0.763(0.097)</td>
<td>-45.214(16.412)</td>
<td>0.757(0.074)</td>
<td>0.446(0.062)</td>
<td>0.317(0.041)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG</td>
<td>1643.492(144.740)</td>
<td>0.500(0.000)</td>
<td>-50.310(16.743)</td>
<td>0.579(0.075)</td>
<td>0.449(0.061)</td>
<td>0.319(0.041)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MI1.8</td>
<td>95.919(53.317)</td>
<td>0.469(0.063)</td>
<td>-45.524(16.489)</td>
<td>0.577(0.075)</td>
<td>0.448(0.062)</td>
<td>0.318(0.040)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MI2.8</td>
<td>103.614(57.700)</td>
<td>0.476(0.062)</td>
<td>-45.744(16.550)</td>
<td>0.577(0.074)</td>
<td>0.448(0.061)</td>
<td>0.319(0.040)</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>MC</td>
<td>1015.501(117.396)</td>
<td>1.500(0.136)</td>
<td>203.799(17.503)</td>
<td>0.288(0.040)</td>
<td>0.212(0.027)</td>
<td>0.149(0.018)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG</td>
<td>4387.649(340.739)</td>
<td>0.500(0.000)</td>
<td>136.526(18.422)</td>
<td>0.325(0.044)</td>
<td>0.242(0.030)</td>
<td>0.174(0.020)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MI1.8</td>
<td>382.907(46.423)</td>
<td>2173.389(257.681)</td>
<td>0.500(0.000)</td>
<td>197.553(17.427)</td>
<td>0.297(0.041)</td>
<td>0.223(0.028)</td>
<td>0.159(0.019)</td>
</tr>
<tr>
<td></td>
<td>MI2.8</td>
<td>464.108(64.943)</td>
<td>2077.739(286.718)</td>
<td>0.500(0.000)</td>
<td>199.878(17.664)</td>
<td>0.299(0.041)</td>
<td>0.222(0.028)</td>
<td>0.162(0.017)</td>
</tr>
<tr>
<td></td>
<td>MI1.8*</td>
<td>382.947(46.658)</td>
<td>2173.238(258.163)</td>
<td>-</td>
<td>197.552(17.428)</td>
<td>0.297(0.041)</td>
<td>0.223(0.028)</td>
<td>0.159(0.019)</td>
</tr>
<tr>
<td>2.5</td>
<td>MC</td>
<td>1003.867(77.651)</td>
<td>2.513(0.165)</td>
<td>550.428(19.452)</td>
<td>0.114(0.022)</td>
<td>0.078(0.013)</td>
<td>0.055(0.009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MG</td>
<td>9804.716(896.551)</td>
<td>0.500(0.000)</td>
<td>318.342(23.092)</td>
<td>0.177(0.029)</td>
<td>0.126(0.018)</td>
<td>0.097(0.011)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MI1.8</td>
<td>573.875(25.735)</td>
<td>5023.835(456.250)</td>
<td>0.500(0.000)</td>
<td>486.069(18.746)</td>
<td>0.137(0.021)</td>
<td>0.104(0.015)</td>
<td>0.076(0.010)</td>
</tr>
<tr>
<td></td>
<td>MI2.8</td>
<td>770.629(35.401)</td>
<td>4185.538(406.434)</td>
<td>0.500(0.000)</td>
<td>512.897(18.693)</td>
<td>0.131(0.022)</td>
<td>0.098(0.014)</td>
<td>0.075(0.008)</td>
</tr>
<tr>
<td></td>
<td>MI1.8*</td>
<td>573.799(25.907)</td>
<td>5023.620(456.576)</td>
<td>-</td>
<td>486.071(18.747)</td>
<td>0.137(0.021)</td>
<td>0.104(0.015)</td>
<td>0.076(0.010)</td>
</tr>
</tbody>
</table>
those values for MC and MI models are comparable. When $\nu = 0.75$, we observe similar results from the cases of $\nu \geq 1.5$. The proposed model has larger maximum loglikelihood value than MG and it has similar performance in prediction to MC. We suspect similar results for $n - \frac{1}{2} < \nu < n$, $n \in \mathbb{N}$. These results suggest that the proposed approach performs significantly better than MG for smooth data, and it has similar performance in prediction to MC.

As we mentioned before, the proposed approach involves a trade-off between the computational cost and accuracy of inference. According to a preliminary study, the performance of the proposed model improves significantly by increasing $n_c$ and $n_p$ for smooth data. Note that MI1.8 and MI1.8* work similarly in terms of maximum likelihood estimation and prediction.

### 3.2. Simulation II

We now use the following compactly supported covariance functions using great circle distance, $\theta \in [0, \pi R]$, which are defined by

\[
K_1(\theta; \alpha, \tau) = \left(1 + \frac{\tau}{\alpha}\right)^{\frac{\tau}{2}} \left(1 - \frac{\theta}{\alpha}\right)^{\frac{\tau}{2}} , \quad \alpha \leq \pi R; \quad \tau \geq 4,
\]

\[
K_2(\theta; \alpha, \tau) = \left(1 + \frac{\tau}{\alpha} + \frac{\tau^2 - 1}{3} \left(\frac{\theta}{\alpha}\right)^2 \right) \left(1 - \frac{\theta}{\alpha}\right)^{\frac{\tau}{2}} , \quad \alpha \leq \pi R; \quad \tau \geq 6.
\]

where $K_1$ and $K_2$ are $C^2$-Wendland and $C^4$-Wendland covariance functions (Gneiting, 2013), respectively, and for any function $g$, $g(\theta)_+ = \max\{g(\theta), 0\}$. Here, $\alpha$ is a range parameter, and $\tau$ is a shape parameter. Both covariance functions are known to be valid on a sphere with great circle distance (Gneiting, 2013). We let $\alpha = \pi R \approx 20,015.09$ (km), $\tau = 14$ for $K_1$, and $\tau = 16$ for $K_2$. We randomly selected 100 locations over the surface of the Earth. Among 100 data points, we used 90 randomly chosen data points for parameter estimation and 10 points for prediction.

For the $C^2$-Wendland case, Table 3 shows that the proposed model performs as well as MC, in terms of maximum loglikelihood and prediction. Note that $C^2$-Wendland function has two derivatives at the origin. Both covariance functions are known to be valid on a sphere with great circle distance (Gneiting, 2013). We let $\alpha = \pi R \approx 20,015.09$ (km), $\tau = 14$ for $K_1$, and $\tau = 16$ for $K_2$. We randomly selected 100 locations over the surface of the Earth. Among 100 data points, we used 90 randomly chosen data points for parameter estimation and 10 points for prediction.

For the $C^2$-Wendland case, Table 3 shows that the proposed model performs as well as MC, in terms of maximum loglikelihood and prediction. Note that $C^2$-Wendland function has two derivatives at the origin. Both covariance functions are known to be valid on a sphere with great circle distance (Gneiting, 2013). We let $\alpha = \pi R \approx 20,015.09$ (km), $\tau = 14$ for $K_1$, and $\tau = 16$ for $K_2$. We randomly selected 100 locations over the surface of the Earth. Among 100 data points, we used 90 randomly chosen data points for parameter estimation and 10 points for prediction.

### 3.3. Simulation III

We used the nonstationary Matérn covariance model that allows for spatially varying local dependence through $\nu$ (Stein, 2005; Paciorek and Schervish, 2006; Kleiber and Nychka, 2012). We let $\nu$ to be different over the sea and the land. We used chordal distance as the distance metric and set $\sigma^2 = 1$, $\alpha = 1000$ (km), $\nu_{\text{sea}} = 1.5$, and $\nu_{\text{land}} = 0.5$. We randomly selected 1000 locations over the surface of the Earth similarly to Section 3.1, and they were used for parameter estimation. For prediction, we also randomly selected 100 locations over four distinct regions (25 locations for each region). One land region near the US states of Colorado is situated between latitudes $38^\circ$ N and $39^\circ$ N, and longitudes $105^\circ$ W and $106^\circ$ W. The other land region near Germany is situated between latitudes $50^\circ$ N and $51^\circ$ N, and longitudes $10^\circ$ E and $11^\circ$ E. Two sea regions are located in South Pacific Ocean.
Table 3
(Simulation Example II) Sample means and standard deviations of parameter estimates, maximum loglikelihood values, and prediction measures for each model (100 cases). True covariance functions are the $C^2$-Wendland function with great circle distance ($\alpha = 20015.09$ and $\tau = 14$), and the $C^4$-Wendland function with great circle distance ($\alpha = 20015.09$ and $\tau = 16$).

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\epsilon}$ or $\hat{\epsilon}_1/\hat{\epsilon}_2$</th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\nu}$</th>
<th>Max.loglik</th>
<th>RMSPE</th>
<th>MAE</th>
<th>CRPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C^2$</td>
<td>MC</td>
<td>1343.405 (359.062)</td>
<td>1.641 (0.309)</td>
<td>73.103 (7.242)</td>
<td>0.190 (0.080)</td>
<td>0.141 (0.053)</td>
<td>0.100 (0.039)</td>
</tr>
<tr>
<td></td>
<td>MG</td>
<td>8074.816 (1564.762)</td>
<td>0.500 (0.000)</td>
<td>51.208 (8.934)</td>
<td>0.234 (0.091)</td>
<td>0.173 (0.063)</td>
<td>0.128 (0.043)</td>
</tr>
<tr>
<td></td>
<td>MI3.6</td>
<td>785.728 (130.844)</td>
<td>0.498 (0.010)</td>
<td>71.195 (7.145)</td>
<td>0.192 (0.080)</td>
<td>0.142 (0.050)</td>
<td>0.100 (0.036)</td>
</tr>
<tr>
<td>$C^4$</td>
<td>MC</td>
<td>1100.667 (194.205)</td>
<td>2.824 (0.451)</td>
<td>137.071 (8.320)</td>
<td>0.078 (0.045)</td>
<td>0.052 (0.023)</td>
<td>0.037 (0.016)</td>
</tr>
<tr>
<td></td>
<td>MG</td>
<td>14315.879 (3752.576)</td>
<td>0.500 (0.000)</td>
<td>75.501 (11.592)</td>
<td>0.130 (0.071)</td>
<td>0.093 (0.042)</td>
<td>0.075 (0.028)</td>
</tr>
<tr>
<td></td>
<td>MI3.6</td>
<td>1127.589 (87.973)</td>
<td>5496.539 (1378.868)</td>
<td>0.498 (0.003)</td>
<td>119.530 (9.012)</td>
<td>0.087 (0.052)</td>
<td>0.059 (0.024)</td>
</tr>
<tr>
<td></td>
<td>MI13.6</td>
<td>736.812 (155.274)</td>
<td>2952.331 (944.473)</td>
<td>0.500 (0.000)</td>
<td>134.284 (8.243)</td>
<td>0.081 (0.049)</td>
<td>0.054 (0.024)</td>
</tr>
</tbody>
</table>
Fig. 3. (Simulation Example 3) One realization of sampling locations, estimation sites (○) and prediction regions (△) for Section 3.3. Each prediction region consists of 25 locations.

(latitudes 13° S and 14° S, and longitudes 170° W and 171° W near Samoa island) and Indian Ocean (latitudes 9° S and 10° S, and longitudes 70° E and 71° E) (see Fig. 3).

We fitted a nonstationary version of the proposed model (call it, NMI), as discussed in Section 2.4. We let $\epsilon$ be different over the sea and the land. Table 4 shows that the NMI2.8 outperforms all stationary covariance models (MC, MG, and MI2.8) in terms of maximum loglikelihood values significantly. Stationary covariance models lead to inadequate model fit. They are unable to capture the local variability because of the global correlation structure. However, by varying $\epsilon$ of NMI, we are able to capture this behavior well.

For prediction, all covariance models have similar performance regarding RMSPE and MAE, but NMI2.8 has the smallest mean CRPS. Moreover, NMI2.8 outperforms in terms of prediction measures over the sea. On the land, all covariance models show similar prediction performances. This may imply that NMI captures the local variability over the sea and the land well by means of spatially varying $\epsilon$.

3.4. Simulation IV

We used three mean zero Gaussian random fields on the surface of the Earth with non-differentiable model, once-, and twice-differentiable models using great circle distance (i.e. MG, MI, and MII, in our notation). We set $\sigma^2 = 1$, $\alpha = 2000$ (km), $\nu = 0.5$ for all models, $\epsilon = 100$ (km) for MI, and $\epsilon_1 = 200$ and $\epsilon_2 = 100$ (km) for MII. For each data generated, we estimated and performed prediction with non-differentiable, once-, and twice-differentiable models. We randomly selected 100 locations over the surface of the Earth. We used 90 randomly chosen data points for parameter estimation and 10 points for prediction. We repeated this process 100 times.

Table 5 displays sample means and standard deviations of the maximum loglikelihood estimates and prediction measures for each model. As we expected, for the data generated from non-differentiable model, all models fitted have similar performance in terms of maximum loglikelihood values and prediction. For the data generated using the model that is only once differentiable, non-differentiable model performed worse than differentiable models in terms of estimation and prediction. For the data generated from a model twice differentiable, both non-differentiable model and the model once differentiable have smaller maximum loglikelihood values than a model twice differentiable. Note that differentiable models perform significantly better than non-differentiable model, for smooth data.
Table 4
(Simulation Example III) Sample means and standard deviations of parameter estimates, maximum loglikelihood values, and prediction measures for each model (100 cases). True covariance function is the nonstationary Matérn model with chordal distance. We set $\sigma^2 = 1$, $\alpha = 1000$, $\nu_{sea} = 1.5$, and $\nu_{land} = 0.5$. For NMI2.8, $\hat{\epsilon}_{sea/land}$, and for prediction measures, RMSPE$_{sea/land}$, MAE$_{sea/land}$, and CRPS$_{sea/land}$ are displayed.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\hat{\epsilon}$</th>
<th>$\hat{\sigma}^2$</th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\nu}$</th>
<th>Max.loglik</th>
<th>RMSPE</th>
<th>MAE</th>
<th>CRPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC</td>
<td>-</td>
<td>1.215(0.182)</td>
<td>1936.720(601.027)</td>
<td>0.620(0.077)</td>
<td>131.484(39.501)</td>
<td>0.191(0.136)</td>
<td>0.162(0.113)</td>
<td>0.140(0.057)</td>
</tr>
<tr>
<td>MG</td>
<td>-</td>
<td>1.219(0.266)</td>
<td>3111.124(764.362)</td>
<td>0.498(0.011)</td>
<td>123.538(36.583)</td>
<td>0.202(0.137)</td>
<td>0.172(0.114)</td>
<td>0.146(0.059)</td>
</tr>
<tr>
<td>MI2.8</td>
<td>37.030(30.337)</td>
<td>1.183(0.189)</td>
<td>2644.205(512.350)</td>
<td>0.500(0.000)</td>
<td>131.512(38.445)</td>
<td>0.195(0.136)</td>
<td>0.165(0.113)</td>
<td>0.142(0.059)</td>
</tr>
<tr>
<td>NMI2.8</td>
<td>502.330(31.042)</td>
<td>0.306(0.627)</td>
<td>1434.010(178.692)</td>
<td>0.497(0.015)</td>
<td>419.501(32.903)</td>
<td>0.166(0.123)</td>
<td>0.143(0.101)</td>
<td>0.103(0.069)</td>
</tr>
</tbody>
</table>

Table 5
(Simulation Example IV) Sample means and standard deviations of parameter estimates, maximum loglikelihood values, and prediction measures for each model (100 cases). True covariance functions are MG($\sigma^2 = 1$, $\alpha = 2000$, $\nu = 0.5$), MI1.8($\sigma^2 = 1$, $\alpha = 2000$, $\nu = 0.5$, $\epsilon = 100$), and MII1.4($\sigma^2 = 1$, $\alpha = 2000$, $\nu = 0.5$, $\epsilon_1 = 200$, $\epsilon_2 = 100$). For MII 1.4, $\hat{\epsilon}_1/\hat{\epsilon}_2$ is displayed.

<table>
<thead>
<tr>
<th>True model</th>
<th>Model</th>
<th>$\hat{\epsilon}$</th>
<th>$\hat{\alpha}$</th>
<th>$\hat{\nu}$</th>
<th>Max.loglik</th>
<th>RMSPE</th>
<th>MAE</th>
<th>CRPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MG</td>
<td>MG</td>
<td>2635.740(1887.045)</td>
<td>0.456(0.076)</td>
<td>$-18.582(6.774)$</td>
<td>0.654(0.179)</td>
<td>0.531(0.153)</td>
<td>0.375(0.100)</td>
<td></td>
</tr>
<tr>
<td>MI1.8</td>
<td>MG</td>
<td>2033.525(435.981)</td>
<td>0.500(0.002)</td>
<td>$-18.756(6.718)$</td>
<td>0.654(0.178)</td>
<td>0.530(0.151)</td>
<td>0.375(0.099)</td>
<td></td>
</tr>
<tr>
<td>MI1.4</td>
<td>MG</td>
<td>204.168(436.111)</td>
<td>0.499(0.005)</td>
<td>$-18.756(6.718)$</td>
<td>0.654(0.178)</td>
<td>0.530(0.151)</td>
<td>0.375(0.099)</td>
<td></td>
</tr>
<tr>
<td>MI1.8</td>
<td>MI1.8</td>
<td>2631.445(960.919)</td>
<td>0.484(0.050)</td>
<td>$-14.048(7.083)$</td>
<td>0.634(0.159)</td>
<td>0.502(0.136)</td>
<td>0.357(0.092)</td>
<td></td>
</tr>
<tr>
<td>MI1.4</td>
<td>MI1.8</td>
<td>2187.339(1065.171)</td>
<td>0.464(0.074)</td>
<td>$-12.867(7.197)$</td>
<td>0.633(0.155)</td>
<td>0.501(0.131)</td>
<td>0.356(0.089)</td>
<td></td>
</tr>
<tr>
<td>MII1.4</td>
<td>MII1.4</td>
<td>3012.419(4834.405)</td>
<td>0.425(0.148)</td>
<td>$-13.215(7.282)$</td>
<td>0.633(0.154)</td>
<td>0.503(0.130)</td>
<td>0.357(0.088)</td>
<td></td>
</tr>
<tr>
<td>MI1.8</td>
<td>MII1.4</td>
<td>2932.159(1887.045)</td>
<td>0.495(0.030)</td>
<td>$-9.481(7.143)$</td>
<td>0.565(0.146)</td>
<td>0.446(0.116)</td>
<td>0.317(0.078)</td>
<td></td>
</tr>
<tr>
<td>MI1.4</td>
<td>MII1.4</td>
<td>2099.711(1091.506)</td>
<td>0.497(0.026)</td>
<td>$-7.160(7.164)$</td>
<td>0.567(0.146)</td>
<td>0.446(0.120)</td>
<td>0.318(0.079)</td>
<td></td>
</tr>
<tr>
<td>MII1.4</td>
<td>MII1.4</td>
<td>2042.154(687.643)</td>
<td>0.480(0.050)</td>
<td>$-6.859(7.177)$</td>
<td>0.569(0.145)</td>
<td>0.448(0.119)</td>
<td>0.318(0.077)</td>
<td></td>
</tr>
</tbody>
</table>
4. Application to climate model output

4.1. Data and the mean structure

We apply the covariance models to sea level pressure data on a global scale. The data is model outputs from the Beijing Climate Center Climate System Model (http://forecast.bcccsm.cma.gov.cn/). It is a part of the fifth phase of the Coupled Model Intercomparison Project (CMIP5). For more information on CMIP5 archive, see Taylor et al. (2012). We use Boreal summer averages (averaged over June to August, JJA, then averaged over 1971–2012). The output values are given on regular grids, and there are 128 longitude points and 64 latitude points; the size of the data is 8192. The unit of the sea level pressure is a Pascal (Pa). For variance stabilization, we took a square root transform of the data, and it is displayed in Fig. 4.

Since we are focusing on modeling the covariance structure of data, we subtracted the mean structure of data and made it close to mean zero. That is, we filtered out mean structure of data by using spherical harmonics, as in Jun and Stein (2008); we regressed the square root of JJA means with \( Y^m_k \{ \sin(L), l \} \) for \( k = 6 \),

\[
f(L, l) = \sum_{k=0}^{6} \sum_{m=-k}^{k} f^m_k Y^m_k \{ \sin(L), l \},
\]

where \( L \) and \( l \) denote latitude and longitude, respectively. Fig. 4 shows the estimated mean structure of data and corresponding residuals.

4.2. Comparison of covariance models

We compared the following covariance models considered so far, MC, MG, Ml(x,y), MI(x,y), and NMl(x,y), for various \( x \) and \( y \) values. Moreover, we considered the nonstationary Matérn covariance model (call it, NMC). For NMC, we let \( \nu \) to be different over the sea and the land. To save computational burden, we randomly selected 1000 locations from the map of residuals. Among 1000 locations, we used 900 randomly chosen locations as training data for parameter estimation, and the remaining 100 data points for validation of prediction. We repeated this process 100 times.

Table 6 provides sample means and standard deviations of parameter estimates for each model. For the models MG and MI, estimates for the range parameter are much larger than that of MC. They appear to be associated with smaller estimates of the smoothness parameter and the estimate of \( \epsilon \). Table 6 also displays sample means and standard deviations of the maximum loglikelihood estimates under each model. The model MII outperforms MI and MG, but it has smaller mean value of maximum loglikelihood than MC. Nonstationary covariance models, NMI and NMC, have larger mean value of maximum loglikelihood than MC. NMC gives somewhat higher loglikelihood value than NMI. For the proposed models, we expect to get better performances in terms of maximum loglikelihood values by increasing \( n_c \) and \( n_p \) further.

Table 6 also shows sample means and standard deviations of prediction measures for each model. As expected, the proposed integrated Matérn models provide smaller RMSPE, MAE, and CRPS values than MG, and their mean value and variability are close to those of MC. Also, MII and NMI perform equally well compared to MC in terms of prediction.

5. Discussion

We have proposed an approach to produce a new class of Matérn-like covariance functions for modeling global data. The resulting covariance functions are positive definite on the surface of a sphere with great circle distance and can deal with smooth processes. There is a straightforward extension of the proposed model for nonstationary data. Through simulation studies and data analysis on a sea level pressure variable, we have shown that new Matérn-like covariance models using great circle distance have substantially better performance in terms of maximum loglikelihood, and prediction
Table 6
(Real data example) Sample means and standard deviations of parameter estimates, maximum loglikelihood values and prediction measures for each model (100 cases). For MII1.4, $\hat{\epsilon}_1/\hat{\epsilon}_2$, for NMI2.6, $\hat{\epsilon}_{sea/land}$, and for NMC, $\hat{\nu}_{sea/land}$ are displayed. For prediction measures, RMSPE$_{sea/land}$, MAE$_{sea/land}$, and CRPS$_{sea/land}$ are displayed.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\hat{\epsilon}$</th>
<th>$\hat{\sigma}^2$</th>
<th>$\hat{\nu}$</th>
<th>Max.loglik</th>
<th>RMSPE</th>
<th>MAE</th>
<th>CRPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MC</td>
<td>-</td>
<td>0.284(0.032)</td>
<td>629.716(101.952)</td>
<td>2.468(0.359)</td>
<td>1347.061(50.962)</td>
<td>0.054(0.017)</td>
<td>0.030(0.006)</td>
</tr>
<tr>
<td>MG</td>
<td>-</td>
<td>0.350(0.016)</td>
<td>5444.964(247.898)</td>
<td>0.500(0.000)</td>
<td>967.223(16.553)</td>
<td>0.087(0.023)</td>
<td>0.053(0.009)</td>
</tr>
<tr>
<td>MII1.8</td>
<td>402.917(33.924)</td>
<td>0.612(0.100)</td>
<td>3948.088(745.099)</td>
<td>0.500(0.000)</td>
<td>1274.958(37.129)</td>
<td>0.058(0.018)</td>
<td>0.032(0.007)</td>
</tr>
<tr>
<td>MII1.4</td>
<td>357.188(84.443)</td>
<td>0.406(0.093)</td>
<td>1815.893(476.291)</td>
<td>0.500(0.000)</td>
<td>1319.975(49.349)</td>
<td>0.059(0.018)</td>
<td>0.032(0.006)</td>
</tr>
<tr>
<td>NMI2.6</td>
<td>636.833(29.948)</td>
<td>0.410(0.081)</td>
<td>2566.736(560.441)</td>
<td>0.500(0.000)</td>
<td>1363.428(30.164)</td>
<td>0.058(0.018)</td>
<td>0.033(0.007)</td>
</tr>
<tr>
<td>NMC</td>
<td>-</td>
<td>0.302(0.038)</td>
<td>1093.846(60.863)</td>
<td>2.351(0.127)</td>
<td>1447.157(37.724)</td>
<td>0.057(0.016)</td>
<td>0.031(0.006)</td>
</tr>
</tbody>
</table>
Fig. 4. (Real data example) Square root of the JJA average of the sea level pressure, its estimated mean structure using spherical harmonics as in (11), and the resulting residuals.

for smooth data when we compare with the Matérn covariance model using great circle distance. In several simulation results and real data analysis presented, unless the spatial range was relatively large, the performance of the proposed approach and Matérn covariance model with chordal distance were similar (except for the nonstationary version of the proposed model). Guinness and Fuentes (2013) report similar findings.
It is evident that larger $n_c$ and $n_p$ offer better approximation to the integration, but this will require substantial increase in computational cost. For computational efficiency with large data and large $n_c$ and $n_p$ values, the proposed model can be fitted with recently developed computational techniques, such as covariance tapering (Furrer et al., 2006; Kaufman et al., 2008) or full scale covariance approximation (Sang and Huang, 2012).

One useful extension for the proposed model is its application to the differential operators approach for nonstationary covariance models on the surface of a sphere (Jun and Stein, 2007, 2008; Jun, 2014). They use chordal distance instead of great circle distance due to mathematical reasons, but adopting the proposed covariance models in their approach will enable them to use great circle distance as a distance metric. Using chordal distance or great circle distance in this case may not make much difference in terms of model fit and spatial prediction in many applications. Nonetheless, as Gneiting (2013) argues, great circle distance is a physically more natural distance metric on the surface of a sphere, and it is worth developing valid covariance models with great circle distance. The proposed model can also be naturally extended to the multivariate case. These are some of the future directions the authors are currently pursuing.

Acknowledgments

Mikyoung Jun’s research was supported by NSF grant DMS-1208421. This publication is based in part on work supported by Award No. KUS-C1-016-04, made by King Abdullah University of Science and Technology (KAUST). The authors acknowledge the modeling groups for making their simulations available for analysis, the Program for Climate Model Diagnosis and Intercomparison (PCMDI) for collecting and archiving the CMIP5 model output, and the World Climate Research Programme (WCRP)’s Working Group on Coupled Modelling (WGCM) for organizing the model data analysis activity. The WCRP CMIP5 multi-model dataset is supported by the Office of Science, US Department of Energy.

References


