VANLANGENBERG, CHRISTOPHER D., Ph.D. Data Generation and Esimation for Axially Symmertic Processes on the Sphere . (2016)
Directed by Dr. Haimeng Zhang. 107 pp.

Data from global networks and satellite sensors have been used to monitor a wide array of processes and variables, such as temperature, precipitation, etc. The modeling and analysis of global data has been extensively studied in the realm of spatial statistics in recent years. In this dissertation, we present our research in the following two areas. In the first project we consider the asymptotics of the popularly used covariance and variogram estimators based on Method of Moments (MOM) for stationary processes on the circle. Although it has been known that such estimators are asymptotically unbiased and consistent when modeling the stationary process on Euclidean spaces, our findings on the circle seem to contradict these results. Specifically, we show that the MOM covariance estimator is biased and the true covariance function may not be identifiable based on this estimator. On the other hand, the MOM variogram estimator is unbiased but inconsistent under the assumption of Gaussianity. Our second research focus is on global data generation. Our proposed parametric models generalize some of existing parametric models to capture the variation across latitudes when modeling the covariance structure of axially symmetric processes on the sphere. We demonstrate that the axially symmetric data on the sphere can be decomposed as Fourier series on circles, where the Fourier random coefficients can be expressed as circularly-symmetric complex random vectors. We develop an algorithm to generate axially symmetric data that follows the given covariance structure. All of the above theories and results are supplemented via simulations.

DATA GENERATION AND ESIMATION FOR AXIALLY SYMMERTIC PROCESSES ON THE SPHERE

by

Christopher D. Vanlangenberg

A Dissertation Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2016

Approved by

Committee Chair

In memory of my fater, Donald Anthony Vanlangenberg

APPROVAL PAGE

This dissertation written by Christopher D. Vanlangenberg has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair ______________________________
Haimeng Zhang

Committee Members ______________________________
Sat Gupta

Scott Richter

Xiaoli Gao

Shan Suthaharan

Date of Acceptance by Committee

Date of Final Oral Examination

ACKNOWLEDGMENTS

Complete later...............!

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

In this chapter we will give a brief introduction to some of the basic concepts in spatial statistics, which are necessary to follow for the rest of this dissertation. More specifically, we will discuss stationarity and intrinsic stationarity, covariance and variogram functions and their properties, mean square continuity and differentiability, spectral representations and spectral densities, complex random processes and Gaussian random vectors, as well as some basic properties related to circulant and block circulant matrices.

1.1 Spatial random field

A random process is a collection of random variables $\{Z(\mathcal{X}) : \mathcal{X} \in \Omega\}$, defined in a common probability space that takes values on a specific domain Ω. Generally, Ω may take a variety of forms as given below.

- $\mathcal{X} \in \Omega = N$: $Z(\mathcal{X})$ is a time series.
- $\mathcal{X} \in \Omega = \mathbb{R}^1$: $Z(\mathcal{X})$ is a random process, commonly referred as a stochastic process.
- $\mathcal{X} \in \Omega = \mathbb{R}^d$: $Z(\mathcal{X})$ is a random field or a spatial process if $d > 1$.
- $\mathcal{X} \in \Omega = \mathbb{S}^2$: $Z(\mathcal{X})$ is a random process on the sphere.
- $\mathcal{X} \in \Omega = \mathbb{R}^d \times R$: $Z(\mathcal{X})$ is a spatio-temporal process that involves both location and time.

We now denote $\{Z(x) : x \in D \subset \mathbb{R}^d\}$ as a real-valued spatial random field in d-dimensional Euclidean space R^d, where x is the location, varying over a fixed domain D. The distribution of $Z(x)$ is characterized by its finite-dimensional distribution function, that is, the distribution function of the random vector $Z(\underset{\sim}{\mathcal{X}}) = (Z(\mathcal{X}_1), \ldots, Z(\mathcal{X}_n))$ given by

$$F(h_1, \ldots, h_n) = P(Z(x_1) \leq h_1, \ldots, Z(x_n) \leq h_n), \tag{1.1}$$

for any n and any sequence of locations $(x_1, x_2, \ldots, x_n)$ and $h_1, \ldots, h_n \in \mathbb{R}$.

Stationarity and Isotropy

A spatial random field $Z(x)$ is said to be strictly stationary, if for any $n,\ x_1, \ldots, x_n \in \mathbb{R}^d$, $h_1, \ldots, h_n \in \mathbb{R}$ and $x \in \mathbb{R}^d$, $Z(x)$ is invariant under translation, that is,

$$P(Z(x_1 + x) \leq h_1, \ldots, Z(x_n + x) \leq h_n) = P(Z(x_1) \leq h_1, \ldots, Z(x_n) \leq h_n) \tag{1.2}$$

Strict stationarity is normally too strong a condition as it involves the distribution of the random field. Another commonly used but weaker assumption is the weak stationarity. More specifically, A random process $Z(x)$ is weakly stationary if

$$\begin{aligned} E(Z(x)) &= \mu \\ E(Z^2(x)) &< \infty \\ C(h) &= Cov(Z(x), Z(x+h)) \end{aligned} \tag{1.3}$$

In other words a random process $Z(x)$ is weakly stationary (or simply stationarity throughout the rest of this dissertation) if it has constant mean and finite second moment as well as its (auto-)covariance function $C(h)$ solely depends on the spatial distance of two locations. Further, a strictly stationary random field with finite second moment is weakly stationary, but weak stationarity does not imply strict stationarity unless $Z(x)$ is a Gaussian random field, under which both stationarities are equivalent, as the finite-dimensional distribution of a Gaussian random field is multivariate normal and, which is uniquely determined by the first and second moments.

The covariance function $C(h)$ of a stationary process $Z(x)$ on $\mathbb{R}^d$ has the following properties.

(i) $C(0) \geq 0$;

(ii) $C(h) = C(-h)$;

(iii) $|C(h)| \leq C(0)$;

(iv) If $C_1(h), C_2(h), \ldots, C_n(h)$ are valid covariance functions, then each of the following functions $C(h)$ is also a valid covariance function.

   (a) $C(h) = a_1C_1(h) + a_2C_2(h)$, $\forall a_1, a_2 \geq 0$;

   (b) $C(h) = C_1(h)C_2(h)$;

   (c) $\lim_{n\to\infty} C_n(h) = C(h), \ \forall h \in \mathbb{R}^d$.

A function $C(\cdot)$ on $\mathbb{R}^d$ is non-negative definite if and only if

$$\sum_{i,j=1}^{N} a_i a_j C(x_i - x_j) \geq 0, \tag{1.4}$$

for any integer N, any constants $a_1, a_2, \ldots, a_N$, and any locations $x_1, x_2, \ldots, x_N \in \mathbb{R}^d$.

A valid covariance must be positive definite. On the other hand, given a positive definite function, one can always define a family $Z(x), x \in D$ of zero-mean Gaussian random process with the given function as its covariance function.

A weakly stationary process with a covariance function $C(||h||)$ which is free from direction is called isotropy. The random field, $Z(x)$, on $\mathbb{R}^d$ is strictly isotropy if the joint distributions are invariant under all rigid motions. *i.e.,* for any orthogonal $d \times d$ matrix H and any $x \in \mathbb{R}^d$

$$P(Z(Hx_1 + x) \leq h_1, \ldots, Z(Hx_n + x) \leq h_n) = P(Z(x_1) \leq h_1, \ldots, Z(x_n) \leq h_n) \tag{1.5}$$

Isotropy assumes that it is not required to distinguish one direction from another for the random field $Z(x)$. When describing the correlation between random fileds at two locations , that is, the correlation of $Z(\mathcal{X})$ at any two locations is the same as long as these two points *.......complete...........*

Variogram function is an alternative to the covariance function proposed by Matheron (1973). It is defined as the variance of the difference between random fields at two locations, that is

$$2\gamma(h) = Var(Z(\mathcal{X}+h) - Z(\mathcal{X})). \tag{1.6}$$

Here $\gamma(h)$ is called the semivariogram. If the variogram function solely depends on the distance of the two locations, then the process with finite constant mean is said to be intrinsically stationary. If $Z(x)$ is further assumed to be stationary with covariance function $C(h)$, then $\gamma(h) = C(0) - C(h)$. Intrinsic stationarity is defined in terms of variogram and it is more general than (weak) stationarity that is defined in terms of covariance. Clearly, when $C(h)$ is known, we can obtain $\gamma(h)$ but the reverse is not true. For example we consider a linear semivariogram function given below,

$$\gamma(h) = \begin{cases} a^2 + \sigma^2 h & h > 0 \\ 0 & otherwise \end{cases}$$

when $\gamma(h) \to \infty$ as $h \to \infty$ thus the process with the above semivariogram is not weakly stationary and $C(h)$ does not exist.

Parallel to the positive definiteness for the covariance function, the variogram is conditionally negative definite, that is,

$$\sum_{i,j=1}^{N} a_i a_j 2\gamma(x_i - x_j) \leq 0, \tag{1.7}$$

for any integer N, any constants $a_1, a_2, \ldots, a_N$ with $\sum_{i=1}^N a_i = 0$, and any locations $x_1, x_2, \ldots, x_N \in \mathbb{R}^d$.

Mean square continuity

For a sequence of random variables $X_1, X_2, \ldots$ and a random variable X defined on a common probability space, define $X_n \xrightarrow{L^2} X$ if $E(X^2) < \infty$ and $E(X_n - X)^2 \to 0$ as $n \to \infty$. We then say $\{X_n\}$ converges in L^2 to X if there exists such a X.

There is no simple relationship between $C(h)$ and the smoothness of $Z(x)$. Suppose $Z(x)$ is a random field on $\mathbb{R}^d$, then $Z(x)$ is mean square continuous at x if

$$\lim_{h \to 0} E(Z(x+h) - Z(x))^2 = 0.$$

If $Z(x)$ is stationary and $C(\cdot)$ is its covariance function then $E(Z(x+h) - Z(x))^2 = 2(C(0) - C(h))$. Therefore $Z(x)$ is mean square continuous if and only if $C(\cdot)$ is continuous at the origin.

Spectral representation of a random field

Suppose $\omega_1, \ldots, \omega_n \in \mathbb{R}^d$ and let $Z_1, \ldots, Z_n$ be mean zero complex-valued random variables with $E(Z_i \bar{Z}_j) = 0$ ($\bar{a}$ represents the conjugate of the complex number a), $i \neq j$ and $E|Z_i|^2 = f_i$. Then the random sum

$$Z(x) = \sum_{k=1}^{n} Z_k e^{i\omega_k^T x}, \quad x \in \mathbb{R}^d \tag{1.8}$$

is a weakly stationary complex random field in $\mathbb{R}^d$ with possibly complex-valued covariance function $C(x) = \sum_{k=1}^{n} f_k e^{i\omega_k^T x}$.

In spatial statistics, sometimes it is more convenient to use complex valued random functions, rather than real valued random functions. We say, $Z(x) = U(x) + iV(x)$ is a complex random field if $U(x), V(x)$ are real random fields. If $U(x), V(x)$ are stationary so does $Z(x)$. The covariance function can be defined as,

$$C(h) = cov(Z(x+h), \overline{Z(x)}), \quad C(-h) = \overline{C(h)}.$$

For any complex constants $c_1, \ldots, c_n$, and any locations $x_1, x_2, \ldots, x_n$,

$$\sum_{i,j=1}^{n} c_i \bar{c}_j C(x_i - x_j) \geq 0 \tag{1.9}$$

Further, if we consider the integral as a limit in L^2 of the above random sum, then the covariance function can be represented as,

$$C(x) = \int_{\mathbb{R}^d} e^{i\omega^T x} F(d\omega) \tag{1.10}$$

where F is the so-called spectral distribution. Here is a more general result from Bochner,

Theorem 1.1 (Bochner's Theorem).

A complex valued covariance function $C(\cdot)$ on $\mathbb{R}^d$ for a weakly stationary mean square *continuous complex-valued random field on* $\mathbb{R}^d$ *if and only if it can be represented as* (1.10) *, where* F *is a positive measure.*

If F has a density (spectral density, denoted by f with respect to Lebesgue measure, (*i.e.* if such a f exists) we can use the inversion formula to obtain f

$$f(\omega) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\omega^T x} C(x) dx \tag{1.11}$$

Spectral densities

Here we provide some examples of isotropic covariance functions and their corresponding spectral densities.

(i) Rational Functions that are even, non-negative and integrable the corresponding covariance functions can be expressed in terms of elementary functions. For example if $f(\omega) = \phi(\alpha^2 + \omega^2)^{-1}$, then $C(h) = \pi\phi\alpha^{-1}e^{-\alpha|h|}$ (obtained by contour integration).

(ii) Gaussian are the most commonly used covariance function for a smooth process on $\mathbb{R}$ where the covariance function is given by $C(h) = ce^{-\alpha h^2}$ and the corresponding spectral density is $f(\omega) = \frac{1}{2\sqrt{\pi\alpha}} ce^{\frac{-\omega^2}{4\alpha}}$.

(iii) $Mat\acute{e}rn$ class has more practical use and more frequently used in spatial statistics. The spectral density of the form $f(\omega) = \frac{1}{\phi(\alpha^2+\omega^2)^{\nu+1/2}}$ where $\phi, \nu, \alpha > 0$ and the corresponding covariance function given by,

$$C(h) = \frac{\pi^{1/2}\phi}{2^{\nu-1}\Gamma(\nu+1/2)\alpha^{2\nu}} (\alpha|h|)^{\nu} Y_{\nu}(\alpha|t|) \tag{1.12}$$

where Y_ν is the modified Bessel function, the larger the ν smoother the Y. Further, Y will be m times square differentiable iff $\nu > m$. When ν is in the form of $m + 1/2$ with m a non negative integer. The spectral density is rational and the covariance function is in the form of $e^{-\alpha|h|}\cdot$ polynomial($|h|$) for example, when $\nu = \frac{1}{2}$ $C(h)$ corresponds to exponential model and $\nu = \frac{3}{2}$ is transformation of exponential family of order 2.

$$\begin{aligned} \nu = 1/2 \quad &: \quad C(h) = \pi\phi\alpha^{-1}e^{-\alpha|h|} \\ \nu = 3/2 \quad &: \quad C(h) = \frac{1}{2}\pi\phi\alpha^{-3}e^{-\alpha|h|}(1 + \alpha|h|) \end{aligned}$$

1.2 Circularly-symmetric Gaussian random vectors

Let $\underset{\sim}{Z} = (Z_1, Z_2, \ldots, Z_n)^T$, where $Z_j = (Z_j^{Re}, Z_j^{Im})^T$ and $j = 1, 2, \ldots, n$ be a zero mean $2n$ complex random vector of dimension $2n$. Then its covariance matrix K_Z and the pseudo-covariance matrix M_Z are defined as follow.

$$K_{\underset{\sim}{Z}} = E[\underset{\sim}{Z}\underset{\sim}{Z}^*] \tag{1.13}$$

$$M_{\underset{\sim}{Z}} = E[\underset{\sim}{Z}\underset{\sim}{Z}^T] \tag{1.14}$$

where $\underset{\sim}{Z}^* = \bar{\underset{\sim}{Z}}^T$ is the conjugate transpose of $\underset{\sim}{Z}$.

Generally, to characterize the relationship of a complex random vector, one needs both covariance and pseudo-covariance matrices. First note that a complex random variable $Z = Z^{Re} + iZ^{Im}$ is (complex) Gaussian, if Z^{Re}, Z^{Im} both are real and they are jointly Gaussian. Now we consider a vector $\underset{\sim}{Z} = (Z_1, Z_2)^T$ where $Z_1 = Z_1^{Re} + iZ_1^{Im}$ and $Z_2 = Z_1^*$ ($Z_2^{Re} = Z_1^{Re}, Z_2^{Im} = -Z_1^{Im}$). The four real and imaginary parts of $\underset{\sim}{Z}$ are jointly Gaussian (each follows $N(0, 1/2)$) (so $\underset{\sim}{Z}$ is complex Gaussian).

The covariance matrices defined by (1.13) is given by

$$M_Z = E \begin{bmatrix} Z_1^2 & Z_1 Z_1^* \\ Z_1 Z_1^* & Z_1^2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

and the pseudo-covariance matrices defined by (1.14) is given by

$$K_Z = E \begin{bmatrix} Z_1 Z_1^* & Z_1^2 \\ Z_1^2 & Z_1 Z_1^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

It is easy to note that $E[Z_1^2] = E[Z_1^{Re} Z_1^{Re} - Z_1^{Im} Z_1^{Im}] = 1/2 - 1/2 = 0$. If both Z_1 and Z_2 are real, then covariance and pseudo-covariance matrices are the same, *i.e.*, $M_Z \equiv K_Z$

The covariance matrix of real $2n$ random vector $\underset{\sim}{Z} = (\underset{\sim}{Z}^{Re}, \underset{\sim}{Z}^{Im})^T$, where $\underset{\sim}{Z}^{Re} = (Z_1^{Re}, Z_1^{Re}, \ldots, Z_1^{Re})$ and $\underset{\sim}{Z}^{Im} = (Z_1^{Im}, Z_2^{Im}, \ldots, Z_n^{Im})$ can be determined by both $K_{\underset{\sim}{Z}}$ and $M_{\underset{\sim}{Z}}$ given as follow.

$$
\begin{aligned}
E[\underset{\sim}{Z}^{Re}\underset{\sim}{Z}^{Re}] &= \frac{1}{2}Re(K_{\underset{\sim}{Z}} + M_{\underset{\sim}{Z}}),\\
E[\underset{\sim}{Z}^{Im}\underset{\sim}{Z}^{Im}] &= \frac{1}{2}Re(K_{\underset{\sim}{Z}} - M_{\underset{\sim}{Z}}),\\
E[\underset{\sim}{Z}^{Re}\underset{\sim}{Z}^{Im}] &= \frac{1}{2}Im(-K_{\underset{\sim}{Z}} + M_{\underset{\sim}{Z}}),\\
E[\underset{\sim}{Z}^{Im}\underset{\sim}{Z}^{Re}] &= \frac{1}{2}Im(K_{\underset{\sim}{Z}} + M_{\underset{\sim}{Z}})
\end{aligned} \tag{1.15}
$$

We can get the covariance of $\underset{\sim}{Z} = (\underset{\sim}{Z}^{Re}, \underset{\sim}{Z}^{Im})^T$ as follows,

$$
\begin{aligned}
Cov(\underset{\sim}{Z}) &= E(\underset{\sim}{Z}\underset{\sim}{Z}^T)\\
&= \begin{pmatrix} E[\underset{\sim}{Z}^{Re}\underset{\sim}{Z}^{Re}] & E[\underset{\sim}{Z}^{Re}\underset{\sim}{Z}^{Im}] \\ E[\underset{\sim}{Z}^{Im}\underset{\sim}{Z}^{Re}] & E[\underset{\sim}{Z}^{Im}\underset{\sim}{Z}^{Im}] \end{pmatrix}
\end{aligned}
$$

Now we introduce circularly-symmetric random variables and vectors. A complex random variable Z is circularly-symmetric if both Z and $e^{i\phi}Z$ have the same probability distribution for all real ϕ. Since $E[e^{i\phi}Z] = e^{i\phi}E[Z]$, any circularly-symmetric complex random variable must have $E[Z] = 0$, in other words its mean must be zero.

For a circularly-symmetric complex random vector, we have the following theorem [Gal08].

Theorem 1.2 (Gallager, 2008)**.**

Let $\underset{\sim}{Z}$ be a zero mean Gaussian random vector then $M_{\underset{\sim}{Z}} = 0$ if and only if $\underset{\sim}{Z}$ is circularly-symmetric.

1.3 Circulant matrix

A square matrix $A_{n\times n}$ is a circulant matrix if the elements of each row (except first row) has the previous row shifted by one place to the right.

$$A = circ[a_o, a_1, \cdots, a_{n-1}] \quad = \begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{bmatrix}. \qquad (1.16)$$

The eigenvalues of A are given by

$$\begin{aligned} \lambda_l &= \sum_{k=0}^{n-1} a_k e^{i2lk\pi/n} \\ &= \sum_{k=0}^{n-1} a_k \rho_l^k, \qquad l = 0, 1, 2, \cdots, n-1, \end{aligned}$$

where $\rho_l = e^{i2\pi l/n}$ represents the lth root of 1, and the corresponding (unitary) eigenvector is given by

$$\psi_l = \frac{1}{\sqrt{n}}(1, \rho_l, \rho_l^2, \cdots, \rho_l^{n-1})^T.$$

If matrix A is real symmetric, that is, $a_i = a_{n-i}$, then its eigenvalues are real. More specifically, for even $n = 2N$ the eigenvalues $\lambda_j = \lambda_{n-j}$ or there are either two eigenvalues or none with odd multiplicity, for odd $n = 2N-1$ the eigenvalue λ_0 equal to any λ_j for $1 \leq j \leq N-1$ or λ_0 occurs with odd multiplicity. A square matrix B is Hermitian, if and only if $B^* = B$ where B^* is the complex conjugate. If B is real then $B^* = B^T$. According to [Tee05] Hermitian matrices has a full set of orthogonal eigenvectors with corresponding real eigenvalues.

Block circulant matrices

The idea of a block circulant matrix was first proposed by [Mui20]. A matrix $B_{np\times np}$ is a block-circulant matrix if it has the following form,

$$B = bcirc[A_o, A_1, \cdots, A_{n-1}] = \begin{bmatrix} A_0 & A_1 & A_2 & \cdots & A_{n-1} \\ A_{n-1} & A_0 & A_1 & \cdots & A_{n-2} \\ A_{n-2} & A_{n-1} & A_0 & \cdots & A_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_1 & A_2 & A_3 & \cdots & A_0 \end{bmatrix} \tag{1.17}$$

where A_j are $(p \times p)$ sub-matrices of complex or real valued elements. [DMG83] proposed some methodologies to find the inverse of B. Let M be a block-permutation matrix

$$M = \begin{bmatrix} 0 & I_p & 0 & \cdots & 0 \\ 0 & 0 & I_p & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_p \\ I_p & 0 & 0 & \cdots & 0 \end{bmatrix},$$

where I_p is $p \times p$ identity matrix and B can be defined as follows,

$$B = \sum_{k=0}^{n-1} A_k M^k.$$

Define M^0 as $(np \times np)$ identity matrix and the eigenvalues of M given by ρ_l, the eigenmatrix of M can be given by $Q_{np\times np} = \{\utilde{\psi_0}, \utilde{\psi_1}, \ldots, \utilde{\psi_{n-1}}\}$. From [Tra73] it can be shown that $Q^{-1} = Q^*/n$ where Q^* is the conjugate transpose of Q now we can write,

$$M = QDQ^{-1} = \frac{QDQ^*}{n}$$

where D is a diagonal matrix and the diagonal elements $D_i \quad i = 0, 1, \ldots, n-1$ are the discrete Fourier transform of the blocks A_j,

$$D_i = \sum_{k=0}^{n-1} A_k e^{i2lk\pi/n}.$$

That is the inverse of matrix B takes the following form,

$$B^{-1} = Q \cdot \begin{pmatrix} D_0^{-1} & 0 & \cdots & 0 \\ 0 & D_1^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_n{-1} \end{pmatrix} \cdot Q^{-1}.$$

The eigenmatrix Q is solely depending on the dimension of B and the eigenvalues of B (ρ_l's), in other words, B is not depending on the blocks (A_j's), *i.e.* for any block diagonal matrix $D_{np\times np}$, QDQ^{-1} is a block circulant matrix, and immediately follows that the inverse of the matrix B is also a block circulant matrix.

When $A_{j_{1\times 1}}$, $B = A$, $D_i^{-1} = \lambda^{-1}$, and the eigenmatrix has a dimension of $n \times n$ then

$$A^{-1} = Q\Lambda^{-1}Q^T \quad \text{where } \Lambda = \{\lambda_0, \ldots, \lambda_{n-1}\}$$

When A is real symmetric Q is also real and symmetric and $Q^{-1} = Q^T$.

*We need to add in the special case when A_j's are symmetric (*by Tee, add citation*)*

CHAPTER II

LITERATURE REVIEW

2.1 Spatial Data

What does it mean by spatial data? In general, spatial data or in other words geospatial data is information about a physical object or a measurement that can be represented by numerical values in a geographic coordinate system. Spatial data appeared to be in the form of maps in 1686 and spatial modeling did not start until 1907 ([Cre93]). There are many questions that geoscientists and engineers are interested about spatial data. Many questions naturally arise such as how to model a spatial process and then use these models to make predictions on the process at unobserved locations. There are many challenges when modeling spatial data; every point (location observed) is a random variable and only one observation/measurement is available. However, the number of unknowns to estimate are quite large compared to the available data, which is definitely a high-dimensional problem. As an example, if data were observed at 10 locations, one is estimating the variance-covariance matrix to characterize the spatial dependency for future predictions, then there will have 55 unknown entities in the variance-covariance matrix to be estimated. We will discuss about some basic properties of geospatial data by exploring some popular data sets in the literature.

Since 1978 Microwave Sounding Units (MSU) measure radiation emitted by the earth's atmosphere from NOAA polar orbiting satellites. The different channels of the MSU measure different frequencies of radiation proportional to the temperature of broad vertical layers of the atmosphere. Tropospheric and lower stratospheric temperature data are collected by NOAA's TIROS-N polar-orbiting satellites and adjusted for time-dependent biases by the Global Hydrology and Climate Center at the University of Alabama in Huntsville (UAH)[1]. More information about how the data is been processed can be found in [CSB00]. Satellites do not measure temperature directly but measure radiances in various wavelength bands and then mathematically inverted to obtain the actual temperature.

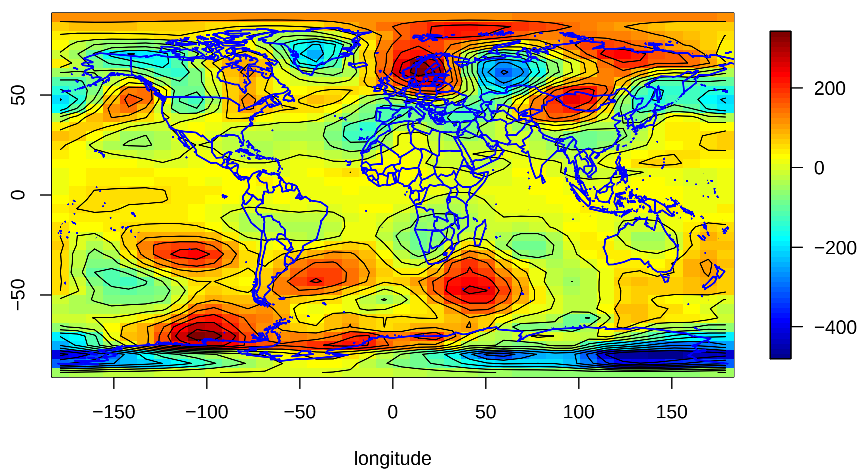
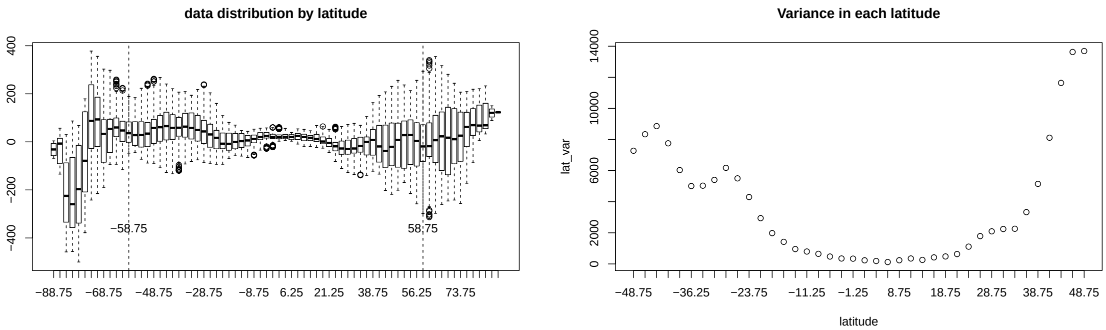


Figure 1. MSU data observed (without removing any spatial trends) in August 2002 : resolution 2.5° latitude × 2.5° longitude (10368 gridded points).

The MSU data were observed at 2.5°latitude × 2.5°longitude with total number of data observations of size 10368.

[1]`https://www.ncdc.noaa.gov/temp-and-precip/msu/overview`



(a) distribution at each latitude (b) variance at each latitude

Figure 2. MSU data distribution at each latitude (data between $60°S$ and $60°N$ were considered)

Level 3 Total Ozone Mapping Spectrometer (TOMS) data [2] is another popluar global data set discussed literature which has more than 20000 sapatial points (gridded points) ([Ste07], [CJ08], [JS08]). There were some missing values in this data set. [Ste07] used the average of 8 neighboring locations to replace the missing values. They used spherical harmonics with associated Legendre polynomials of up to 78 covariates to remove the spatial trends to study axially symmetry (discussed in chapter 4) of the global data.

[2]`http://disc.sci.gsfc.nasa.gov/data/datapool/TOMS`

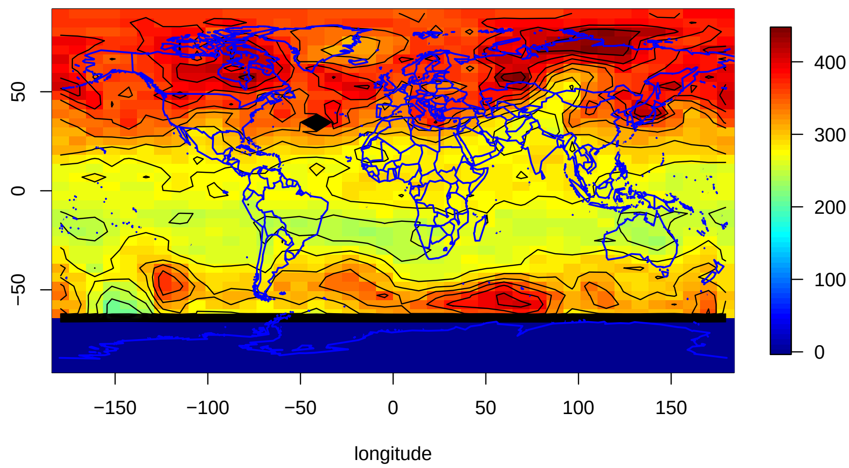
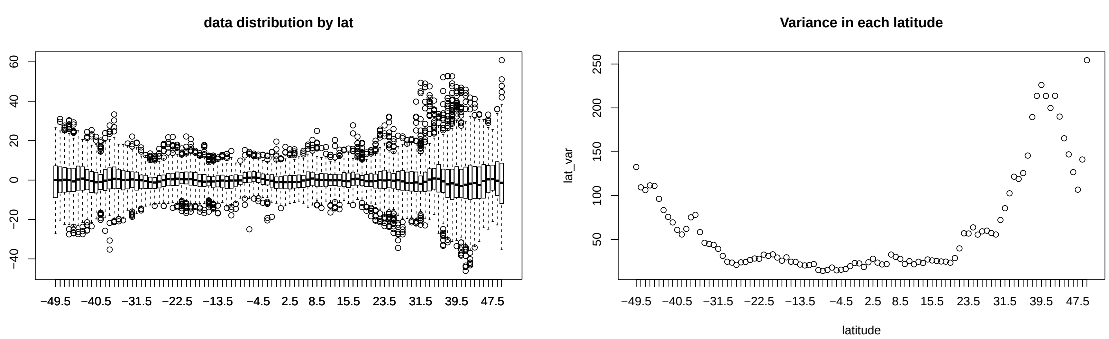


Figure 3. TOMS data: resolution 1°latitude $\times 1.25°$ longitude in May, 1-6 1990. The instrument used backscattered sunlight, therefore measurements were not available south of $73°S$ during this week.



(a) distribution at each latitude

(b) variance at each latitude

Figure 4. TOMS data distribution at each latitude (data between $50°S$ and $50°N$ were considered)

Both MSU and TOMS data demonstrate strong variation when it is closer to Earth's poles. This shows the complexity of geospatial data. In addition, collecting geospatial data could be very expensive and time consuming. Next we discuss about the challenges when modeling spatial data.

2.2 Literature review

There have been extensive statistical research on methodologies and techniques developed under the Euclidean space R^d. These approaches that are valid in R^d have been applied to analyze global-scale data in recent years, due to global networks and satellite sensors that have been used to monitor a wide array of global-scale processes and variables. Here are some commonly used covariance models that are valid on R^3.

Family	C(h)	Parameters
$Mat\acute{e}rn$	$\frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\frac{h}{\phi})^\nu Y_\nu(\frac{h}{\phi})$	ν, σ^2, ϕ
Spherical	$\sigma^2(1-\frac{3h}{2\phi}+\frac{1}{2}(\frac{h}{\phi})^3)I_{0\leq h\leq\phi}$	ϕ, σ^2
Exponential	$\sigma^2 exp\{-(h/\phi)\}$	ϕ, σ^2
Gaussian	$\sigma^2 exp\{-(h/\phi)^2\}$	ϕ, σ^2

Table 1. Commonly used covariance models in $\mathbb{R}^3$

However, this can have unforeseen impacts, such as making use of models that are valid in R^d but in fact might not be valid under spherical coordinate systems. [HZR11] have investigated some of commonly used covariance models that are valid in R^d, and they pointed out that many of those are actually invalid on the sphere. In particular they indicated (later proved by [Gne13]) that the commonly used $Mat\acute{e}rn$ covariance model is not valid if the smoothness parameter (ν) is greater than 0.5, when modeling the homogeneous processes on the Earth.

The main emphasis of this dissertation is on random processes on a unit sphere. In particular, we focus on covariance models and estimation in modeling the global data. Note that the assumption of homogeneity on the sphere requires that the mean of the random process is constant and the covariance function of the process at two locations depends only on the spherical distance. This assumption is difficult to evaluate and often deemed unrealistic in practice. Several approaches on modeling non-homogeneity have been proposed in literature. For example, [Ste07] argued that Total Ozone Mapping Spectrometer (TOMS) data varies strongly with latitudes and homogeneous models are not suitable. Monte Carlo Markov Chain (MCMC) is another approach to model non-stationary covariance models on a sphere. [LRL11] analyzed global temperature data with a non-stationary model defined on a sphere using Gaussian Markov Random Fields (GMRF) and Stochastic Partial Differential Equations (SPDE). Further [BL11] constructed a class of stochastic field models using SPDEs and non stationary covariance models were obtained by spatially varying the parameters in the SPDEs, where they claim that the method is more efficient than standard MCMC procedures. Furthermore, aerosol depth (AOD) from Multi-angle Imaging Spectrometer (MISR), Sea Surface Temperature (SST) from RRMM Microwave Imager (TMI) are some other examples for anisotropy global data on a sphere.

The analysis and modeling of axially symmetric data on the sphere has received increasing attention in literature in recent years. It was first introduced in [Jon63], where the covariance of random processes between two spatial points depends on the longitudes only through their difference between those points. This assumption is more plausible and reasonable when modeling spatial data. For example, geophysical processes or variables such as temperature, moisture often exist homogeneity on longitudes rather than latitudes. [Ste07] used spherical harmonics to model (TOMS) data that exhibit an axial symmetry. [JS08] applied first-order differential operators to an isotropic process to draw conclusions about the local properties of axially symmetric spatio-temporal processes. proposed a flexible class of parametric covariance models to capture the nonstationarity using the differential operators. [HS12] investigates the properties and theory of different forms of axially symmetric processes on the sphere. [Li13] used convolution methods with $Mat\acute{e}rn$-type kernel functions to capture the non-stationarity of random fields on a sphere. [HZR12] developed a new and simplified representation for a valid axially symmetric process as well as explore the construction of parametric models for axially symmetric processes.

The computational cost for modeling and analyzing the axially symmetric data is very expensive. As we will see later in Chapter 4, the covariance function for axially symmetric processes requires triple summations, which one is to estimate $\mathcal{O}(n^3)$ parameters. Although the covariance structure given by [HZR12] might potentially reduce the number of parameters to be estimated in the order of $\mathcal{O}(n^2)$, the large data sets from global sensors and satellites often add much more computational cost. [Ste07] used 170 parameters from axially symmetric covariance structure to model TOMS data but was still not able to capture the global dependency. [CJ08] used more than 396 parameters when they model the global data. Therefore, it is necessary to develop practically useful parametric models with easily interpretable parameters.

Statistical simulations have been one of the critical components in statistical research. Through simulations, the researcher can explore how a proposed statistical model/method behaves in the simulated and reproducible data that mimic the real applications. For axially symmetric processes, however, it seems in literature that there is lacking of the simulation method to generate global data that follow the axially symmetric covariance structure. Note that, as we will see later, the spectral representation of the process on the sphere is a summation of Legendre polynomials, which is distinct from its planar counterpart as represented by an integration of Bessel functions in R^d. This distinction could be understood through group representation theory, which possibly lies on the compactness of a sphere. Therefore, the estimation methods proposed based on R^d should be reexamined for validity. All these areas are the basis for an exciting new line of research we are currently pursuing.

2.3 The outline of this dissertation

In Chapter 3 we explore some of the properties of commonly used covariance and variogram estimators on the circle based on Method of Moments (MOM). In contrasting to the results given in time series and the Euclidean space, the MOM covariance estimator is biased and the true covariance function is non identifiable based on the MOM estimator. On the other hand, the MOM variogram estimator is unbiased, but it can be shown to be inconsistent. In Chapter 4 we first introduce the random process on the sphere. We then discuss the homogeneous process and the spectral representation for its covariance function on the sphere. Our main focus on this chapter is the axially symmetric process and its covariance function representation through the discrete fourier transform. The parametric models for characterizing such processes are also discussed. In particular, we extend the models given in [HZR12] and provide some graphical properties of those models. These generalized models will be fully implemented in Chapter 5 for axially symmetric data generation. In Chapter 5, we explore the result given in [HZR12] to implement an algorithm for axially symmetric data generation. In particular, we observe that our proposed method on data generation perform better based on cross variogram. Finally, Chapter 6 gives a summary of this research and provides further research directions.

CHAPTER III

ASYMPTOTICS OF ESTIMATORS ON A CIRCLE

3.1 Introduction

For a stationary time series process, the covariance and variogram function estimators have been commonly used in literature. More specifically, let $X(t)$ be a stationary process with unknown constant mean μ and covariance function $C(h), h \geq 0$. Let $\gamma(h) = C(0) - C(h)$ be the variogram function. For simplicity, we assume $\{X(t_k), k = 1, 2, \ldots, n\}$ is a sequence of random variables observed on gridded locations $\{t_k = (k-1)\delta, k \geq 1\}$ in R^1 with $\delta = t_k - t_{k-1} > 0$ be the fixed interval length. First, note that $\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X(t_k)$ is an unbiased estimator of μ, and further under the certain ergodic condition, for example, the covariance function converges to zero as the lag $h \to \infty$, one has

$$var(\bar{X}) \to 0, \qquad \text{as } n \to \infty,$$

implying the consistency of $\bar{X}$ when estimating μ.

The covariance and variogram function estimators based on the method of moments (MOM) are given by

$$\begin{aligned} \hat{C}(h) &= \frac{1}{n-h}\sum_{j=1}^{n-h}(X(t_j+h) - \bar{X})(X(t_j) - \bar{X}), \qquad h = 1, 2, \cdots, n-1. \\ 2\hat{\gamma}(h) &= \frac{1}{n-h}\sum_{j=1}^{n-h}(X(t_j+h) - X(t_j))^2, \qquad h = 1, 2, \cdots, n-1, \end{aligned}$$

respectively. It has been shown, for example, Cressie(1993), that both the variogram and covariance MOM estimators are biased while the bias for the covariance MOM estimator is of the rate of $O(1/n)$ but the bias for the variogram MOM estimator is smaller. Furthermore, the asymptotic joint Gaussianity of the estimators $\{\hat{C}(h)\}$ and of $2\hat{\gamma}(h)$ has also been established under the same ergodic conditions (*i.e.*, conditions that ensure the dependence in the process dies off sufficiently quickly as lag distance increases). Finally, for fixed h, both the variances and covariances of $\{\hat{C}(h)\}$ and of $2\hat{\gamma}(h)$ can be found (for example, in Fuller (1976) and Cressie (1985)) are of $O(1/n)$, which demonstrates the consistency of both estimators. Parallel results can also be obtained for the covariance and variogram MOM estimators in R^d ([Cre93]).

There are two distinct asymptotics in spatial statistics: increasing domain asymptotics, where more data are collected by increasing the domain, and fixed-domain or infill asymptotics, where more data are collected by sampling more densely in a fixed domain. Asymptotic properties of estimators are quite different under the two asymptotics. The above asymptotics in the time series and R^d belongs to the first one. There have been an extensive discussion of the infill asymptotics in literature, for example, Zhang (2004) showed that for the popularly used Matérn covariance model, one cannot correctly distinguish between two Matérn covariances with probability one no matter how many sample data are observed in a fixed region. Consequently, not all covariance parameters in Matérn models are consistently estimable. However, there seems lacking of discussions about the asymptotics on the circle and sphere. In particular, with the increasing interest in the study of the global data on the sphere, it is necessary that the infill asymptotics of the existing estimators need to be examined. As the first step for the study of the asymptotics of covariance and variogram estimators on the sphere, we consider their asymptotics on the circle.

This chapter is organized as follow. We first introduce the random processes on the circle and then give the spectral representation of the covariance and variogram functions for stationary processes. Under the assumption of stationary Gaussian process on the circle, we show that the unbiased estimator $\bar{X}$ is not consistent when estimating μ. Then, we demonstrate that the covariance function estimator based on MOM is biased, and may not be estimable, while the variogram MOM estimator is unbiased but inconsistent. Finally we present some simulation results.

Random process on a circle

Let $X(t)$ be the random process on the unit circle S^1. If $X(t)$ is further assumed to be with finite second moment and continuity in quadratic mean, then $X(t)$ can be represented in a Fourier series which is convergent in quadratic mean ([DR76]),

$$X(t) = A_0 + \sum_{n=1}^{\infty}(A_n \cos(nt) + B_n \sin(nt)), \qquad t \in S^1,$$

where

$$A_0 = \frac{1}{2\pi}\int_S X(t)dt, \quad A_n = \frac{1}{\pi}\int_S X(t)\cos(nt)dt, \quad B_n = \frac{1}{\pi}\int_S X(t)\sin(nt)dt.$$

Let $s, t \in S^1$. The covariance function $C(s,t)$ of the process $X(t)$ on the given locations s and t is given below

$$C(s,t) = cov(X(s), X(t)).$$

Now we assume the underlying process $X(t)$ is stationary on the circle, that is, $E(X(t)) = \mu$ unknown, and its covariance function solely depends on the angular distance θ

$$C(\theta) = cov(X(t+\theta), X(t)), \qquad \theta \in [0, \pi].$$

Under the assumption of stationarity, we have

$$cov(A_n, A_m) = cov(B_n, B_m) = a_n \delta(n, m), \qquad \text{and} \qquad cov(A_n, B_m) = 0, \quad \text{for } n \geq 0, m > 0,$$

with $a_n \geq 0$, under which the covariance function $C(\theta)$ can be written as the following spectral representation

$$C(\theta) = a_0 + \sum_{n=1}^{\infty} a_n \cos(n\theta), \qquad \text{for } \theta \in [0, \pi].$$

By the orthogonality of $\{\cos(n\theta), n = 0, 1, 2, \cdots,\}$ on $\theta \in [0, \pi]$, we have

$$a_0 = \frac{1}{\pi} \int_0^{\pi} C(\theta) d\theta, \qquad a_n = \frac{2}{\pi} \int_0^{\pi} C(\theta) \cos(n\theta) d\theta, \quad n \geq 1.$$

If a random process $X(t)$ is intrinsically stationary, one has $E(X(t)) = \mu$, an unknown constant, as well as the variogram function depends only on the angular distance θ, given by

$$\gamma(\theta) = var(X(t+\theta) - X(t)), \qquad t \in S^1.$$

Note that if $X(t)$ is stationary, then

$$\gamma(\theta) = C(0) - C(\theta).$$

Equivalently, $\gamma(\theta)$ has the following spectral representation

$$\gamma(\theta) = \sum_{n=1}^{\infty} a_n(1 - \cos(n\theta)).$$

Mean and covariance estimation on the circle

We now consider the estimation on the unknown mean μ and covariance function $C(\theta)$. Let $\{X(t_k), k = 1, 2, \cdots, n\}$ be a collection of gridded observations on the unit circle, with $t_k = (k-1) * 2\pi/n, k = 1, 2, \cdots, n$. For simplicity, let $n = 2N$ be an even number. Denote $\underset{\sim}{X} = (X_1, X_2, \cdots, X_n)^T$ as the observed random vector. When the underlying process $X(t)$ is stationary on the unit circle, the variance-covariance matrix of $\underset{\sim}{X}$ is given by

$$\Sigma = \begin{pmatrix} C(0) & \cdots & C((N-1)\delta) & C(\pi) & C((N-1)\delta) & \cdots & C(\delta) \\ C(\delta) & \cdots & C((N-2)\delta) & C((N-1)\delta) & C(\pi) & \cdots & C(2\delta) \\ C(2\delta) & \cdots & C((N-3)\delta) & C((N-2)\delta) & C((N-1)\delta) & \cdots & C(3\delta) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ C(\delta) & \cdots & C(\pi) & C((N-1)\delta) & C((N-2)\delta) & \cdots & C(0) \end{pmatrix},$$

is a symmetric circulant matrix with elements $C(0), C(\delta), C(2\delta), \cdots, C((N-1)\delta), C(\pi), C((N-1)\delta), \cdots, C(\delta)$ where $\delta = 2\pi/n$. Therefore the sample mean

$$\bar{X} = \frac{1}{n}\mathbf{1}_n^T \underset{\sim}{X}$$

is an unbiased estimator of μ with the variance given by

$$\begin{aligned} var(\bar{X}) &= cov\left(\frac{1}{n}\mathbf{1}_n^T \underset{\sim}{X}, \frac{1}{n}\mathbf{1}_n^T \underset{\sim}{X}\right) \\ &= \frac{1}{n^2}\mathbf{1}_n^T \Sigma \mathbf{1}_n \\ &= \frac{1}{n}\left(C(0) + C(\pi) + 2\sum_{m=1}^{N-1} C(m2\pi/n)\right) \end{aligned}$$

If we assume that $C(\theta)$ is a continuous function on $[0, \pi]$ and note the summation in the last quantity is a trapezoid sum of $C(\theta)$ on the gridded locations within $[0, \pi]$, then,

$$\frac{1}{\pi}\frac{\pi}{2N}\left(C(0) + \sum_{m=1}^{N-1} C(m2\pi/n) + C(\pi)\right) \to \frac{1}{\pi}\int_0^{\pi} C(\theta)d\theta = a_0,$$

as $n \to \infty$. That is, $var(\bar{X}) \to a_0$ as $n \to \infty$. Therefore, we have the following proposition.

Proposition 3.1. *The sample mean $\bar{X}$ is an unbiased estimator of μ with the asymptotic variance of a_0. If $a_0 > 0$ and $X(t)$ is further assumed to be Gaussian, then $\bar{X}$ is NOT a consistent estimator of μ.*

Proof. It is only necessary to prove the second part. If $X(t)$ is Gaussian, then $\bar{X} \sim N(\mu, var(\bar{X})) \Rightarrow Z = \frac{\bar{X}-\mu}{\sqrt{var(\bar{X})}} \sim N(0,1)$. First, for a fixed $\varepsilon_0 > 0$ and $\varepsilon_0 < a_0$, there exists K, such that for all $n > K$, we have

$$|var(\bar{X}) - a_0| < \varepsilon_0 \Rightarrow a_0 - \varepsilon_0 < var(\bar{X}) < a_0 + \varepsilon_0.$$

Now for each fixed $\varepsilon > 0$ and all $n > K$,

$$\begin{aligned} P\left(|\bar{X} - \mu| > \varepsilon\right) &= P\left(\frac{|\bar{X} - \mu|}{\sqrt{var(\bar{X})}} > \frac{\varepsilon}{\sqrt{var(\bar{X})}}\right) \\ &\geq P\left(|Z| > \frac{\varepsilon}{\sqrt{a_0 - \varepsilon_0}}\right) > 0. \end{aligned}$$

Hence $\bar{X} \not\to \mu$ in probability. The last inequality above is due to the following.

$$\left\{|Z| > \frac{\varepsilon}{\sqrt{a_0 - \varepsilon_0}}\right\} \subseteq \left\{|Z| > \frac{\varepsilon}{\sqrt{var(\bar{X})}}\right\}.$$

□

Remark 1. Under the assumption of Gaussianity, Proposition 3.1 indicates that $\bar{X}$ will never be a consistent estimator for μ, which contrasts to the result given in time series and R^d.

Remark 2. Recall that, if $X(t)$ is stationary process on the circle, we have

$$a_0 = var(A_0)$$

and $\mu = E(A_0)$. Therefore, $a_0 = 0 \Rightarrow \mu = 0$, under which $var(\bar{X}) \to 0$ and so $\bar{X}$ is consistent. In other words, if $X(t)$ is a zero mean stationary process on the circle, then $\bar{X}$ is an unbiased and consistent estimator of $\mu = 0$.

Remark 3. If in practice, we have multiple copies of data observations on the circle, we can then estimate a_0 or $var(\bar{X})$ through these copies. More explicitly, suppose that we have *i.i.d.* copies of the random samples on the circle with averages denoted as $\{\bar{X}_i, i = 1, 2, \cdots, m\}$. We then use the method of moments for estimating a_0, which is given as following:

$$\hat{a}_0 = \frac{1}{m-1}\sum_{j=1}^{m}(\bar{X}_j - \bar{\bar{X}})^2,$$

where $\bar{\bar{X}} = \frac{1}{m}\sum_{k=1}^{m}\bar{X}_k$. Under some regularity conditions, one can show that $\hat{a}_0$ is unbiased and consistent estimator of a_0.

Now we consider the MOM estimator of $C(\theta)$, which is given by

$$\hat{C}(\Delta\lambda) = \frac{1}{n}\sum_{i=1}^{n}(X(t_i + \Delta\lambda) - \bar{X})(X(t_i) - \bar{X}), \tag{3.1}$$

where $\Delta\lambda = 0, 2\pi/n, 4\pi/n, \cdots, 2(N-1)\pi/n$.

3.2 Data generation on a circle

To explore the properties of the MOM estimator, we performed a simple simulation study. We consider two covariance fucntions that are valid on a circle, exponential family and power family as given below,

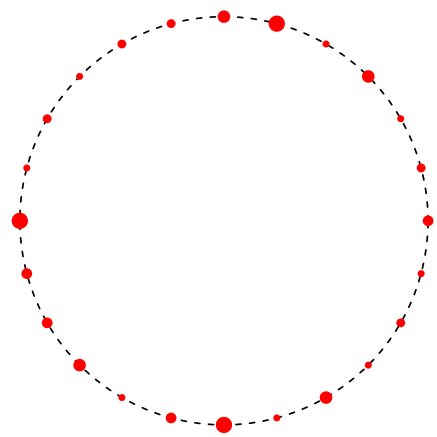


Figure 5. Random process on a circle at 24 points ($\Delta\lambda = 15^0$), the red dots represent the observed values at a given time and each point is associated with a random process of it's own.

$$C(\theta) = C_1 e^{-a|\theta|} \quad a > 0, C_1 > 0 \tag{3.2}$$

$$C(\theta) = c_0 - (|\theta|/a)^{\alpha} \quad a > 0, \alpha \in (0, 2] \tag{3.3}$$

where $\theta = i * \Delta\lambda = \pm i * 2\pi/n, i = 1, 2, \cdots, \lfloor n/2 \rfloor$ and $c_0 \geq \int_0^{\pi} (\theta/a)^{\alpha} \sin\theta d\theta$

Clearly, each location is correlated with other $n-1$ locations and $C(\theta) = C(-\theta)$ the variance-covariance matrix Σ is circulant. We use singular value decomposition (SVD) and obtain the correlated data $\underset{\sim}{X}$ on a circle as follows,

$$\underset{\sim}{X} = \Sigma^{1/2} * Z = Q\Lambda^{1/2}Q^T * Z,$$

where $\Lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ and $Q = \{\psi_1, \psi_2, \cdots, \psi_n\}$ are eigen values and eigen vectors of the circulant matrix respectively and $Z \sim N(\underset{\sim}{0}, 1_n)$.

Covariance estimator comparison

In general the covarince estimator (3.1) on a circle is biased, with a bias of $var(\bar{X})$. In order to make things simple we set $C_1, a = 1$ and when $\alpha = 0.5$ $c_0 \geq \int_0^{\pi} (\theta)^{0.5} \sin\theta d\theta$, from Fresnel intergal it can be shown that $c_0 \geq 2.4353$, for simplicity we set $c_0 = \pi$. Now we compare the covariance estimator (empirical) to it's theoretical covariance given by (3.2) and (3.3). We computed the MOM estimator $\hat{C}(\theta)$ with 48 gridded observations on the circle from 500 simulations.

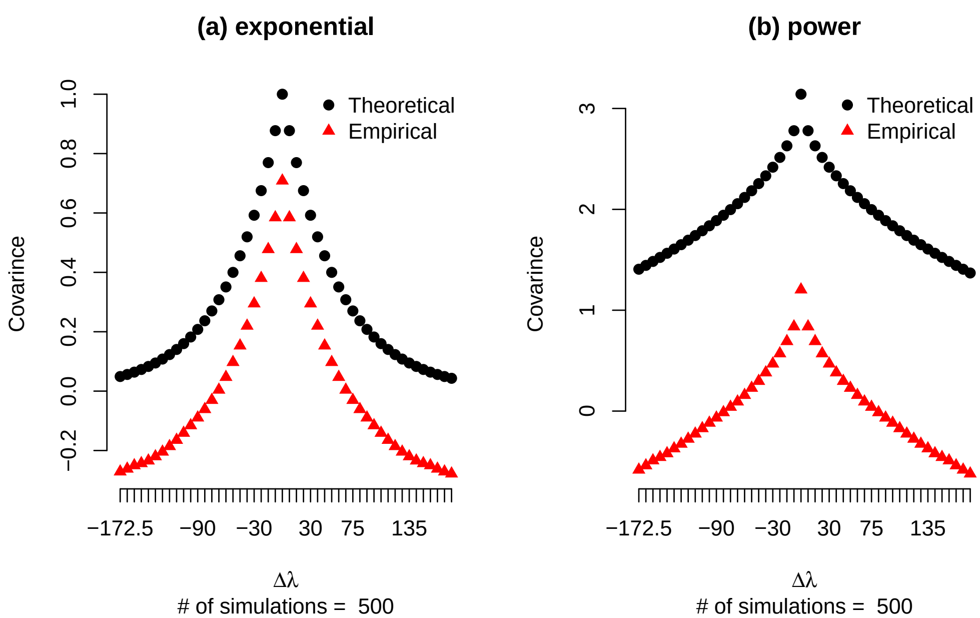


Figure 6. Theoretical and empirical covariance (with bias) comparison on a circle, it is easy to notice the bias in both covariance models

Remark 4. The shift between theoretical and empirical values is equal to a_0 and from $a_0 = \frac{1}{\pi}\int_0^\pi C(\theta)d\theta$ we can obtain

$$\begin{aligned} \text{exponential}: a_0 &= \frac{C_1}{a\pi}(1 - e^{-a\pi}) \\ \text{power}: a_0 &= c_0 - \left(\frac{\pi}{a}\right)^\alpha \frac{1}{\alpha + 1} \end{aligned}$$

Now consider the following covariance function, after subtracting a_0 from $C(\theta)$.

$$D(\theta) = C(\theta) - a_0.$$

If the new covariance function $D(\theta)$ was used to generate the data on a circle then the covariacne estimate converge to the theoretical value. In other words if the process is a zero mean process the covariance estimator given by (3.1) is unbiased $(i.e. Var(\bar{X}) = 0)$ hence we will get a perfect match between theoretical and empirical values.

Remark 5. The covariance estiamtor is biased and the biasness will approach to a_0. When covariance function is unknown the biaseness a_0 is also known and the biasness cannnot to estimated (cannot find the variance of $\bar{X}$) from one circle, however if multiple *i.i.d.* copies on the same circle were available then one can estimate a_0 *i.e.* $\hat{a} = var(\bar{X})$ and substract $\hat{a_0}$ form the MOM estimator as given below,

$$\hat{C}(\Delta\lambda) = \left(\frac{1}{n}\sum_{i=1}^{n}(X(t_i + \Delta\lambda) - \bar{X})(X(t_i) - \bar{X})\right) - \hat{a_0}$$

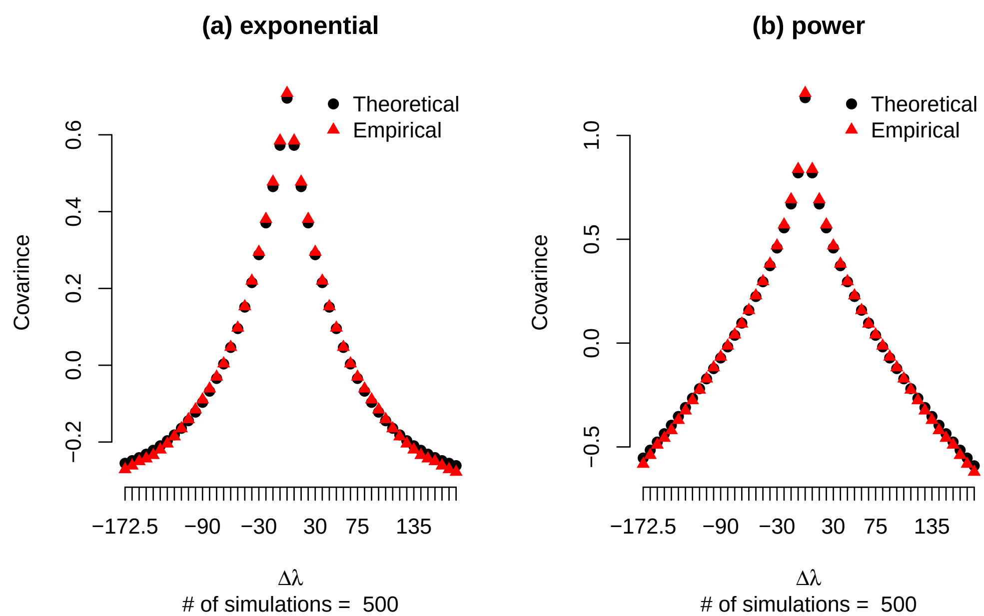


Figure 7. Theoretical and empirical covariance comparison on a circle using the covariance function $D(\theta)$.

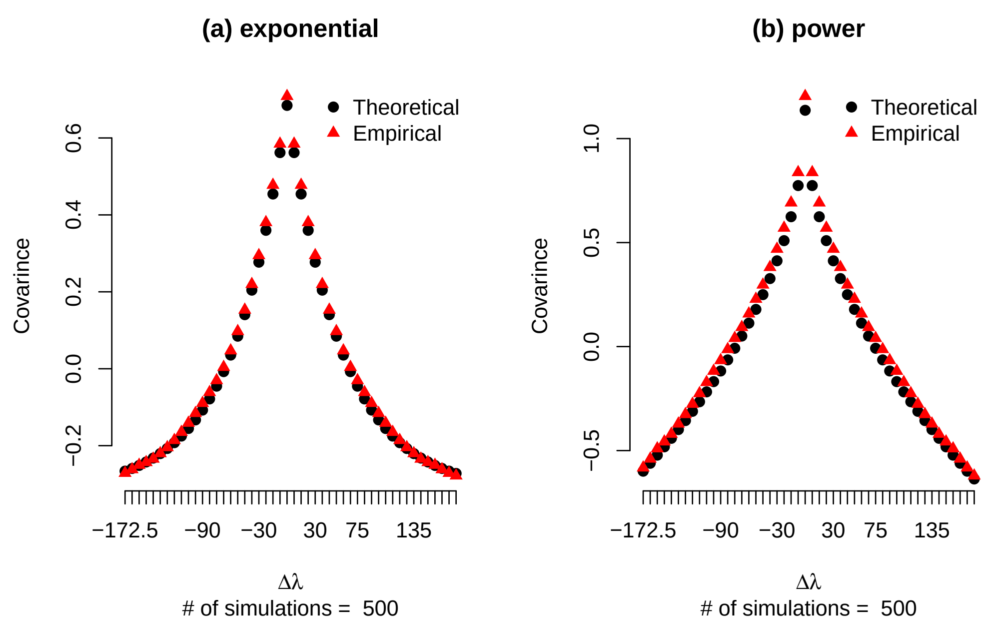


Figure 8. Theoretical and empirical covariance comparison on a circle, after removing $\hat{a_0}$ from the

Now we theoretically calculate the unbiasedness and consistency of the above estimator.

$$\begin{aligned}
E(\hat{C}(\Delta\lambda)) &= \frac{1}{n}\sum_{i=1}^{n} E((X(t_i+\Delta\lambda)-\bar{X})(X(t_i)-\bar{X})) \\
&= \frac{1}{n}\sum_{i=1}^{n} E((X(t_i+\Delta\lambda)-\mu-(\bar{X}-\mu))(X(t_i)-\mu-(\bar{X})-\mu)) \\
&= \frac{1}{n}\sum_{i=1}^{n} cov(X(t_i+\Delta\lambda),X(t_i)) - \frac{1}{n}\sum_{i=1}^{n} E((X(t_i+\Delta\lambda)-\mu)(\bar{X}-\mu)) \\
&\quad -\frac{1}{n}\sum_{i=1}^{n} E((X(t_i)-\mu)(\bar{X}-\mu)) + \frac{1}{n}\sum_{i=1}^{n} E((\bar{X}-\mu)(\bar{X}-\mu)) \\
&= C(\Delta\lambda) - E((\bar{X}-\mu)(\bar{X}-\mu)) - E((\bar{X}-\mu)(\bar{X}-\mu)) \\
&\quad +E((\bar{X}-\mu)(\bar{X}-\mu)) \\
&= C(\Delta\lambda) - var(\bar{X}).
\end{aligned}$$

That is, the MOM estimator $\hat{C}(\Delta\lambda)$ of the covariance function is actually a biased estimator with the shift amount approximately equal to a_0 unless $a_0 = 0$. In other words, if $a_0 = 0$, the MOM estimator $\hat{C}(\Delta\lambda)$ is an unbiased estimator of $C(\theta)$. In summary, we have

Proposition 3.2. *The MOM covariance estimator is a biased estimator of the true covariance function $C(\theta)$, if $a_0 > 0$. However, if the process is a zero mean process then the MOM covariance estimator is unbiased.*

Variogram estimator comparison

In general the variogram estimator in the case of a is unbiased but not consistent. When the random process on a circle is isotropy the semi variogram is given by

$$\gamma(\theta) = C(0) - C(\theta),$$

the theoretical variogram based on exponential and power covariance functions can be given in the following form,

$$\begin{aligned} \text{exponential}: \gamma(\theta) &= C(0) - C(\theta) = C_1(1 - e^{-a|\theta|}) \\ \text{power}: \gamma(\theta) &= C(0) - C(\theta) = (|\theta|/a)^\alpha \end{aligned}$$

We computed the variogram estimator $\hat{\gamma}(\theta)$ with 48 gridded observations on the circle from 500 simulations since there is no bias varigram estimator is a better fit between theoretical and empirical values compared to covariance estimate.

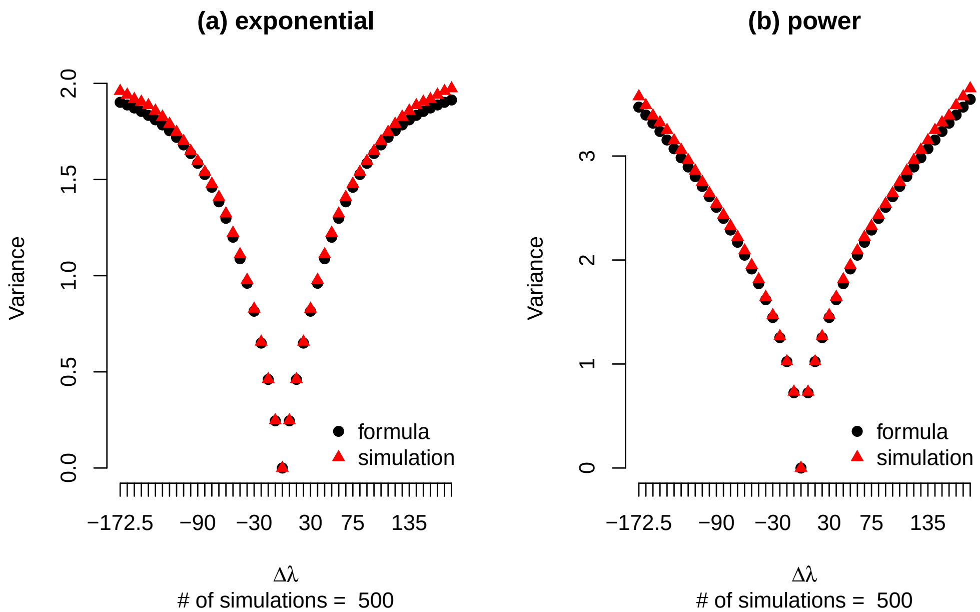


Figure 9. Theoretical and empirical comparison for variogram on a circle

Variogram estimation

In R^n, The variogram estimator generally performs better than the covariance function estimator [Cre93]. Given gridded data observations $\underset{\sim}{X}$, the variogram estimator through Method of Moments is given by

$$\hat{\gamma}(\Delta\lambda) = \frac{1}{2n}\sum_{i=1}^{n}(X(t_i + \Delta\lambda) - X(t_i))^2, \tag{3.4}$$

$$
\begin{aligned}
\text{Then } E(\hat{\gamma}(\Delta\lambda)) &= \frac{1}{2n}\sum_{i=1}^{n} E(X(t_i+\Delta\lambda)-X(t_i))^2 \\
&= \frac{1}{2n}\sum_{i=1}^{n} E((X(t_i+\Delta\lambda)-\mu)-(X(t_i)-\mu))^2 \\
&= \frac{1}{2n}\sum_{i=1}^{n} cov(X(t_i+\Delta\lambda)-X(t_i), X(t_i+\Delta\lambda)-X(t_i)) \\
&= \frac{1}{2n}\sum_{i=1}^{n} (cov(X(t_i+\Delta\lambda), X(t_i+\Delta\lambda)) + cov(X(t_i), X(t_i)) \\
&\quad - 2cov(X(t_i+\Delta\lambda), X(t_i))) \\
&= \frac{1}{2n}\sum_{i=1}^{n} (C(0)+C(0)-2C(\Delta\lambda)) \\
&= C(0)-C(\Delta\lambda) = \gamma(\Delta\lambda).
\end{aligned}
$$

Therefore, $\hat{\gamma}(\Delta\lambda)$ is an unbiased estimator of $\gamma(\Delta\lambda)$.

We first calculate the variance and covariance of the variogram estimator on the circle. Again we consider the equal-distance gridded points on the circle $\{t_i : 1 \leq i \leq n, t_i = (i-1)\times 2\pi/n\}$ and $\underset{\sim}{X} = (X(t_1), X(t_2), \ldots, X(t_n))^T$, being the observed data values. Assume that the random process $X(t)$ is stationary. Note that the Matheron's classical semi-variogram estimator on the circle based on the method of moments can be written as

$$
\hat{\gamma}(\Delta\lambda) = \underset{\sim}{X}^T A(\Delta\lambda)\underset{\sim}{X}.
$$

Here for all $\Delta\lambda$, $A(\Delta\lambda)$ is a circulant matrix, and in particular, $A(0) = 0$. For simplicity, we set $n = 2N$ to be even. First we give an example $n = 6$ to demonstrate the structure of $A(\Delta\lambda)$.

Let $n = 6$. We have four distance angles $\Delta\lambda = 0, \pi/3, 2\pi/3, \pi$. Then each of design matrices $A(\Delta\lambda)$ is given below.

$$
\begin{aligned}
A(0) &= 0 \\
A(\pi/3) &= \frac{1}{12}\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & -1 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ -1 & 0 & 0 & 0 & -1 & 2 \end{pmatrix} = \frac{1}{12} circ(2,-1,0,0,0,-1); \\
A(2\pi/3) &= \frac{1}{12}\begin{pmatrix} 2 & 0 & -1 & 0 & -1 & 0 \\ 0 & 2 & 0 & -1 & 0 & -1 \\ -1 & 0 & 2 & 0 & -1 & 0 \\ 0 & -1 & 0 & 2 & 0 & -1 \\ -1 & 0 & -1 & 0 & 2 & 0 \\ 0 & -1 & 0 & -1 & 0 & 2 \end{pmatrix} = \frac{1}{12} circ(2,0,-1,0,-1,0); \\
A(\pi) &= \frac{1}{12}\begin{pmatrix} 2 & 0 & 0 & -2 & 0 & 0 \\ 0 & 2 & 0 & 0 & -2 & 0 \\ 0 & 0 & 2 & 0 & 0 & -2 \\ -2 & 0 & 0 & 2 & 0 & 0 \\ 0 & -2 & 0 & 0 & 2 & 0 \\ 0 & 0 & -2 & 0 & 0 & 2 \end{pmatrix} = \frac{1}{12} circ(2,0,0,-2,0,0).
\end{aligned}
$$

In general, for $1 \leq m \leq N-1$, and let $\delta = 2\pi/n$ be the common interval length so that $\Delta\lambda = m\delta$. Then we have

$$\begin{aligned}
A(0) &= 0; \\
A(m\delta) &= \frac{1}{2n} circ(2, 0, 0, \ldots, -1, 0, \ldots, -1, 0, \ldots, 0), \\
&\quad \text{where } -1\text{'s are placed at } (m+1)^{th} \text{ and } (n-m+1)^{th} \text{ positions, respectively;} \\
A(N\delta) &= A(\pi) = \frac{1}{2n} circ(2, 0, 0, \ldots, -2, 0, \ldots, 0), \\
&\quad \text{where } -2 \text{ is placed at } (N+1)\text{th position.}
\end{aligned}$$

Obviously, $A(\Delta\lambda) = A(m\delta)$ is a symmetric circulant matrix. From Section 1.3, the eigenvalues of $A(m\delta)$ is then given by

$$\begin{aligned}
\lambda_j^{(A)} &= \frac{1}{2n}(2 - (\exp(j2\pi i/n))^m - (\exp(j2\pi i/n))^{n-m}) \\
&= \frac{1}{2n}(2 - \exp(mj2\pi i/n) - \exp(-mj2\pi i/n)) \\
&= \frac{1}{n}(1 - \cos(jm\lambda)) = \frac{1}{n}(1 - \cos(j\Delta\lambda)), \qquad j = 0, 1, 2, \ldots, n-1.
\end{aligned}$$

for $1 \leq m \leq N-1$, and for $m = N$,

$$\begin{aligned}
\lambda_j^{(A)} &= \frac{1}{2n}(2 - 2(\exp(j2\pi i/n))^N) \\
&= \frac{1}{n}(1 - \cos(j\pi)) \\
&= \frac{1}{n}(1 - \cos(j\Delta\lambda)), \qquad j = 0, 1, \cdots, n-1.
\end{aligned}$$

In addition, from Section 1.3, all circulant matrices can be orthogonally diagonalized using the same orthogonal (Fourier) matrix, denoted as P. Consequently, the trace of the product of circulant matrices = the trace of product of diagonal matrices = the sum of the product of corresponding eigenvalues from those circulant matrices.

Now we consider the distribution of the variogram estimator. First we write the variogram estimator in the following form

$$\begin{aligned}\hat{\gamma}(\Delta\lambda) &= \frac{1}{2n}\sum_{i=1}^{n}(X(t_i+\Delta\lambda)-X(t_i))^2 \\ &= \frac{1}{2n}\sum_{i=1}^{n}((X(t_i+\Delta\lambda)-\mu)-(X(t_i)-\mu))^2.\end{aligned}$$

Therefore,

$$\hat{\gamma}(\Delta\lambda) = (\underset{\sim}{X} - \underset{\sim}{\mathbf{1}}_n\mu)^T A(\Delta\lambda)(\underset{\sim}{X} - \underset{\sim}{\mathbf{1}}_n\mu) \tag{3.5}$$

Note that $A(\Delta\lambda)$ is a circulant matrix with following spectral decomposition

$$A(\Delta\lambda) = P\Lambda^{(A)}P^T,$$

where P is the so-called fourier matrix (orthogonal), solely depending on the dimension of A, and

$$\begin{aligned}\Lambda^{(A)} &= \operatorname{diag}(\lambda_1^{(A)}, \lambda_2^{(A)}, \cdots, \lambda_n^{(A)}), \\ \text{with} \quad \lambda_m^{(A)} &= \frac{1}{n}(1-\cos((m-1)\Delta\lambda)), \qquad m = 1, 2, \cdots, n.\end{aligned}$$

If $\underset{\sim}{X}$ follows a multivariate normal $N(\underset{\sim}{\mathbf{1}}_n\mu, \Sigma)$, then $(\underset{\sim}{X} - \underset{\sim}{\mathbf{1}}_n\mu) \sim N(\underset{\sim}{\mathbf{0}}, \Sigma)$. Note that the variance-covariance matrix Σ is also a circulant matrix, which has the following spectral decomposition.

$$
\begin{aligned}
\Sigma &= P\Lambda^{(\Sigma)}P^T, \\
\text{with} \quad \Lambda^{(\Sigma)} &= \operatorname{diag}(\lambda_1^{(\Sigma)}, \lambda_2^{(\Sigma)}, \cdots, \lambda_n^{(\Sigma)}),
\end{aligned}
$$

$$
\text{where } \lambda_j^{(\Sigma)} = \left(C(0) + 2\sum_{m=1}^{N-1} C(m\delta)\cos((j-1)m\delta) + C(\pi)\cos((j-1)N\delta) \right).
$$

Let $\underset{\sim}{Y} = P^T\left(\underset{\sim}{X} - \underset{\sim}{\mathbf{1}}_n\mu\right)$, then $\underset{\sim}{Y}$ follows a multivariate normal distribution with mean $\underset{\sim}{\mathbf{0}}$ and variance-covariance matrix given by

$$
\begin{aligned}
var(\underset{\sim}{Y}) &= cov(P^T\left(\underset{\sim}{X} - \underset{\sim}{\mathbf{1}}_n\mu\right), P^T\left(\underset{\sim}{X} - \underset{\sim}{\mathbf{1}}_n\mu\right)) \\
&= P^T\Sigma P = P^TP\Lambda^{(\Sigma)}P^TP = \Lambda^{(\Sigma)}.
\end{aligned}
$$

That is, $\underset{\sim}{Y} = (Y_1, Y_2, \cdots, Y_n)^T$ are independent normal random variates with mean 0 and variance $\lambda_j^{(\Sigma)}$.

The variogram estimator is then given by

$$
\begin{aligned}
\hat{\gamma}(\Delta\lambda) &= (\underset{\sim}{X} - \underset{\sim}{\mathbf{1}}_n\mu)^T A(\Delta\lambda)(\underset{\sim}{X} - \underset{\sim}{\mathbf{1}}_n\mu) \\
&= (P(\underset{\sim}{X} - \underset{\sim}{\mathbf{1}}_n\mu))^T\Lambda^{(A)}(P^T(\underset{\sim}{X} - \underset{\sim}{\mathbf{1}}_n\mu)) \\
&= \underset{\sim}{Y}\Lambda^{(A)}\underset{\sim}{Y} = \sum_{m=1}^{n} \lambda_m^{(A)} Y_m^2.
\end{aligned}
$$

Note $\frac{Y_m}{\sqrt{\lambda_m^{(\Sigma)}}} \sim N(0,1)$, and so $\frac{Y_m^2}{\lambda_m^{(\Sigma)}} \sim \chi_1^2$ (or written as $\chi_{1,m}^2$ due to the dependency on m), which implies

$$\hat{\gamma}(\Delta\lambda) = \sum_{m=1}^{n} \lambda_m^{(A)} \lambda_m^{(\Sigma)} \left(\frac{Y_m}{\sqrt{\lambda_m^{(\Sigma)}}} \right)^2 \triangleq \sum_{m=1}^{n} \lambda_m^{(A)} \lambda_m^{(\Sigma)} \chi_{1,m}^2.$$

Here $\chi_{1,1}^2, \chi_{1,2}^2, \cdots, \chi_{1,n}^2$ are *i.i.d.* following χ_1^2 distribution. Hence

$$E(\hat{\gamma}(\Delta\lambda)) = \sum_{m=1}^{n} \lambda_m^{(A)} \lambda_m^{(\Sigma)}, \qquad var(\hat{\gamma}(\Delta\lambda)) = 2 \sum_{m=1}^{n} (\lambda_m^{(A)} \lambda_m^{(\Sigma)})^2$$

After tedious calculations, one can show that

$$E(\hat{\gamma}(\Delta\lambda)) = \sum_{m=1}^{n} \lambda_m^{(A)} \lambda_m^{(\Sigma)} = \gamma(\Delta\lambda),$$

which recovers the result we had early. Next we consider the variance of the variogram estimator under the Gaussian assumption.

Without loss of generality, we assume that $a_1 > 0$ (otherwise, we can always choose some a_m such that $a_m > 0$). First notice that

$$\begin{aligned} \hat{\gamma}(\Delta\lambda) &= \sum_{m=1}^{n} \lambda_m^{(A)} \lambda_m^{(\Sigma)} \chi_{1,m}^2 \\ &= (C(0) - C(\Delta\lambda)) \sum_{m=1}^{n} \frac{\lambda_m^{(A)} \lambda_m^{(\Sigma)}}{C(0) - C(\Delta\lambda)} \chi_{1,m}^2 \\ &\triangleq (C(0) - C(\Delta\lambda)) \sum_{m=1}^{n} C_{n,m} \chi_{1,m}^2, \end{aligned}$$

where $\sum_{m=1}^{n} C_{n,m} = \sum_{m=1}^{n} \frac{\lambda_m^{(A)}\lambda_m^{(\Sigma)}}{C(0)-C(\Delta\lambda)} = 1$ and $C_{n,m} > 0$ since both the matrices A and Σ are positive definite. Hence

$$
\begin{aligned}
var(\hat{\gamma}(\Delta\lambda)) &= (C(0) - C(\Delta\lambda))^2 * 2 * \left(\sum_{m=1}^{n} C_{n,m}^2\right) \\
&\leq 2(C(0) - C(\Delta\lambda))^2 \left(\sum_{m=1}^{n} C_{n,m}\right) = 2(C(0) - C(\Delta\lambda))^2.
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
var(\hat{\gamma}(\Delta\lambda)) &= (C(0) - C(\Delta\lambda))^2 * 2 * \left(\sum_{m=1}^{n} C_{n,m}^2\right) \\
&\geq (C(0) - C(\Delta\lambda))^2 * 2 * C_{n,2}^2.
\end{aligned}
$$

Note that

$$
\begin{aligned}
C_{n,2} &= \frac{1-\cos(\Delta\lambda)}{C(0) - C(\Delta\lambda)} \frac{1}{n} \left(C(0) + 2\sum_{k=1}^{N-1} C(k\delta)\cos(k\delta) + C(\pi)\cos(N\delta) \right) \\
&= \frac{1-\cos(\Delta\lambda)}{C(0) - C(\Delta\lambda)} \frac{1}{\pi} \frac{\pi}{n} \left(C(0) + 2\sum_{k=1}^{N-1} C(k\delta)\cos(k\delta) + C(\pi)\cos(N\delta) \right).
\end{aligned}
$$

Now we consider the limit of $C_{n,2}$ when $n \to \infty$. We want to point out that when $n \to \infty$, we are sampling denser and denser data points over the circle so that we are always maintaining $\Delta\lambda$ to be fixed. A simple approach is to take the sample size n to be doubled so that n tends to infinity while maintaining $\gamma(\Delta\lambda)$ to be estimable. Under this setting, we have

$$
\frac{\pi}{n} \left(C(0) + 2\sum_{k=1}^{N-1} C(k\delta)\cos(k\delta) + C(\pi)\cos(N\delta) \right) \to \int_0^{\pi} C(\theta)\cos(\theta)d\theta, \quad \text{as } n \to \infty,
$$

and so

$$\begin{aligned}
C_{n,2} &\to \frac{1-\cos(\Delta\lambda)}{C(0)-C(\Delta\lambda)}\frac{1}{2}\frac{2}{\pi}\int_0^\pi C(\theta)\cos(\theta)d\theta = \frac{1-\cos(\Delta\lambda)}{C(0)-C(\Delta\lambda)} * \frac{a_1}{2} \\
C_{n,2} &> \frac{1-\cos(\Delta\lambda)}{C(0)-C(\Delta\lambda)} * (\frac{a_1}{2} - \varepsilon_0)
\end{aligned}$$

for a fixed $0 < \varepsilon_0 < \frac{a_1}{2}$ and a large enough sample size n. Consequently,

$$var(\hat{\gamma}(\Delta\lambda)) \quad > \quad 2(a_1/2 - \varepsilon_0)^2(1-\cos(\Delta\lambda))^2.$$

We summary our findings as the following proposition.

Proposition 3.3. *The variogram MOM estimator is finite and asymptotically bounded away from zero.*

From our previous calculation, we have, for each fixed m,

Now we present our main result for the MOM variogram estimator.

Proposition 3.4. *If the underlying process $X(t)$ is assumed to be Gaussian, the MOM variogram estimator is not consistent on the circle.*

Proof. First we consider the consistency of the variogram estimator. To show the following

$$P\left(|\hat{\gamma}(\Delta\lambda) - \gamma(\Delta\lambda)| \geq \varepsilon\right) \to 0,$$

as $n \to \infty$ for fixed $\varepsilon > 0$ and $\Delta\lambda \neq 0$, it is equivalent to show that

$$P\left(\left|\sum_{m=1}^{n} C_{n,m}\chi^2_{1,m} - 1\right| \geq \varepsilon\right) \to 0,$$

as $n \to \infty$ for fixed $\varepsilon > 0$ and $\Delta\lambda \neq 0$. Here $\sum_{m=1}^{n} C_{n,m} = 1, C_{n,m} > 0$ for each fixed n. Note that we also have for each fixed m,

$$0 < C_{n,m} \to \frac{a_m}{2}\frac{1-\cos(m\Delta\lambda)}{C(0)-C(\Delta\lambda)} \equiv b_m.$$

For simplicity, we can assume that $b_2 > 0$ (Otherwise we can pick some $b_m > 0$ for some m fixed). That is

$$C_{n,2} \to b_2 > 0, \qquad \text{as } n \to \infty.$$

Therefore, for fixed $\varepsilon_0 > 0$ and $\varepsilon_0 < b_2$, we choose all $n > N$, such that

$$b_2 - \varepsilon_0 < C_{n,2} < b_2 + \varepsilon_0$$

Therefore, for all $n > N$, (and denote $\chi^2_{1,2} = \chi^2_1$ for simplicity)

$$\sum_{m=1}^{n} C_{n,m}\chi^2_{1,m} \geq C_{n,2}\chi^2_{1,2} > (b_2 - \varepsilon_0)\chi^2_1$$

Hence notice that, for the fixed $\varepsilon > 0$,

$$\left\{(b_2 - \varepsilon_0)\chi^2_1 > 1 + \varepsilon\right\} \subseteq \left\{\sum_{m=1}^{n} C_{n,m}\chi^2_{1,m} > 1 + \varepsilon\right\}$$

Now, for all $n \geq N$,

$$
\begin{aligned}
& P\left(\left|\sum_{m=1}^{n} C_{n,m}\chi^2_{1,m} - 1\right| \geq \varepsilon\right) \\
= \; & P\left(\sum_{m=1}^{n} C_{n,m}\chi^2_{1,m} > 1+\varepsilon \quad \text{or} \quad \sum_{m=1}^{n} C_{n,m}\chi^2_{1,m} < 1-\varepsilon\right) \\
\geq \; & P\left(\sum_{m=1}^{n} C_{n,m}\chi^2_{1,m} > 1+\varepsilon\right) \geq P\left((b_2-\varepsilon_0)\chi^2_1 > 1+\varepsilon\right) \\
= \; & P\left(\chi^2_1 > \frac{1+\varepsilon}{b_2-\varepsilon_0}\right) \nrightarrow 0,
\end{aligned}
$$

since the last term is a fixed positive number. This proves the non-consistency of variogram estimator.

$\square$

CHAPTER IV

PARAMETRIC MODELS ON A SPHERE

4.1 Random process on a sphere

Suppose $X \in \{X(P) : P \in D\}$, defined in a common probability space $P \in S^2$ (unit sphere), where $P = (\lambda, \phi) \in S^2$ with longitude $\lambda \in [-\pi, \pi)$ and latitude $\phi \in [0, \pi]$. Suppose the process is continuous in quadratic mean with respect to the location P and has finite second moment, then the process can be represented by spherical harmonics ([Jon63],[LN97, HZR12]), with the sum converging in the sense of mean squares.

$$X(P) = \sum_{\nu=0}^{\infty} \sum_{m=-\nu}^{\nu} Z_{\nu,m} e^{im\lambda} P_\nu^m(\cos\phi),$$

Here $P_\nu^m(\cdot)$ are normalized associated Legendre polynomials such that their squared integral on $[-1, 1]$ is 1, and $Z_{\nu,m}$ are complex-valued coefficients satisfying

$$Z_{\nu,m} = \int_{S^2} X(P) e^{-im\lambda} P_\nu^m(\cos\phi) dP.$$

Without loss of generality, we suppose that the process $X(P)$ is with zero mean, which implies $E(Z_{\nu,m}) = 0$. Let $P = (\lambda_P, \phi_P)$ and $Q = (\lambda_Q, \phi_Q)$ be two arbitrary locations on the sphere, the covariance function of the process is given by

$$\begin{aligned} R(P,Q) &= \mathrm{E}(X(P)\overline{X(Q)}) \\ &= \sum_{\nu=0}^{\infty}\sum_{\mu=0}^{\infty}\sum_{m=-\nu}^{\nu}\sum_{n=-\mu}^{\mu} \mathrm{E}(Z_{\nu,m}\overline{Z}_{\mu,n}) e^{im\lambda_P} P_\nu^m(\cos\phi_P) e^{-in\lambda_Q} P_\mu^n(\cos\phi_Q), \end{aligned}$$

where $\bar{Z}$ denotes the complex conjugate of Z. Note that the continuity of $X(P)$ on every point P implies that $R(P,Q)$ is continuous at all pairs of (P,Q) [Lea67, page 83].

Homogeneous covariance functions on sphere

Under the assumption of homogeneity (or isotropy), the covariance function of a random process $X(\cdot)$ on S^2 is invariant under the rotations. More specifically, a homogeneous random process on the sphere satisfies

$$\begin{aligned} E(X(P)) &= \mu \quad \text{for any } P \in S^2 \\ Cov(X(P),X(Q)) &= C(\theta_{PQ}) \end{aligned}$$

where θ_{PQ} is the spherical angle between two locations P, Q, which is given by

$$\theta_{PQ} = \arccos\left(\sin(\phi_P)\sin(\phi_Q) + \cos(\phi_P)\cos(\phi_Q)\cos(\lambda_P - \lambda_Q)\right).$$

Parallel to the requirement for a valid covariance function in $\mathbb{R}^d$, a valid covariance function $C(\cdot)$ must be non-negative definite, *i.e.*,

$$\sum_{i,j=1}^{N} a_i a_j C(\theta_{P_iP_j}) \geq 0,$$

for any integer N, any constants $a_1, a_2, \ldots, a_N$, and any locations $P_1, P_2, \ldots, P_N \in S^2$.

According to [Sch42], a real continuous function $C(\theta)$ is a valid covariance function on S^2 if and only if it can be written in the following form

$$C(\theta) = \sum_{k=0}^{\infty} c_k P_k(\cos\theta),$$

where $P_k(\cdot)$ is the Legendre polynomial, and $\forall c_k \geq 0$ and $\sum_k c_k < \infty$. A general result of the above representation on S^d can also be found in [Sch42].

Note the Legendre polynomials $P_k(\cdot)$ are orthogonal in the following sense

$$\int_{-1}^{1} P_n(x)P_m(x)dx = \frac{2}{2n+1}\delta_{nm},$$

with $\delta_{nm} = 1$ if $n = m$ and 0 otherwise. Hence the coefficients c_k can be obtained by

$$c_k = \frac{2k+1}{2}\int_{0}^{\pi} C(\theta)P_k(\cos\theta)d\theta. \quad k = 0, 1, 2, \ldots \tag{4.1}$$

One can directly use the above integral to evaluate the validity of a covariance function on the sphere by checking if c_k is non-negative and $\sum_k c_k < \infty$.

The construction of covariance models is critical for spatial prediction. However, the covariance models that are valid on $\mathbb{R}^d$ may not be valid on the sphere (S^2). For example, [HZR11] evaluated the validity of commonly used covariance that are valid on $\mathbb{R}^d$ and summarized their findings in the following table.

Model	Covariance function	Validity S^2
Spherical	$\left(1-\frac{3\theta}{2a}+\frac{1}{2}\frac{\theta^3}{a^3}\right)\mathbf{1}_{(\theta\leq a)}$	Yes
Stable	$\exp\left\{-\left(\frac{\theta}{a}\right)^\alpha\right\}$	Yes for $\alpha\in(0,1]$
		No for $\alpha\in(1,2]$
Exponential	$\exp\{-\left(\frac{\theta}{a}\right)\}$	Yes
Gaussian	$\exp\left\{-\left(\frac{\theta}{a}\right)^2\right\}$	No
Power*	$c_0-(\theta/a)^\alpha$	Yes for $\alpha\in(0,1]$
		No for $\alpha\in(1,2]$
Radon transform of order 2	$e^{-\theta/a}(1+\theta/a)$	No
Radon transform of order 4	$e^{-\theta/a}(1+\theta/a+\theta^2/3a^2)$	No
Cauchy	$(1+\theta^2/a^2)^{-1}$	No
Hole - effect	$\sin a\theta/\theta$	No

Table 2. Validity of covariance functions on the sphere, $a>0,\theta\in[0,\pi]$. *When $\alpha\in(0,1]$, power model is valid on the sphere for some $c_0\geq\int_0^\pi(\theta/a)^\alpha\sin\theta d\theta$.

Furthermore, they conjectured (later proved by [Gne13]) that the Mat$\acute{e}$rn covariance function is only valid on the sphere when the smoothness parameter $\nu\in(0,1/2]$. Other way of constructing the valid homogeneous covariance function on the sphere is by using the valid covariance function in R^3. Specifically, [Yad83] showed that if

$K(\cdot)$ is valid isotropic covariance function on $\mathbb{R}^3$ then

$$C(\theta) = K(2\sin(\theta/2))$$

is a valid isotropic covariance function on the unit sphere, where θ is the greatest circle distance on the sphere.

Variogram on a sphere

Parallel to the case of circle, if a random process $X(\cdot)$ on a sphere is intrinsically stationary on S^2, then one has $E(X(P)) = \mu$, an unknown constant for all $P \in S^2$ and the variogram function between any two locations $P, Q \in S^2$ depends only on the spherical angle θ_{PQ}

$$Var(X(P) - X(Q)) = 2\gamma(\theta_{PQ}) \quad , \forall P, Q \in S^2$$

The variogram function is conditionally negative definite, that is,

$$\sum_{i,j=1}^{N} a_i a_j 2\gamma(\theta_{P_i P_j}) \leq 0,$$

for any integer N, any constants $a_1, a_2, \ldots, a_N$ with $\sum_i a_i = 0$, and any locations $P_1, P_2, \ldots P_N \in S^2$. Immediately from **??**, for a continuous $2\gamma(\cdot)$ with $\gamma(0) = 0$ the variogram is negative definite if and only if

$$\gamma(\theta) = \sum_{k=0}^{\infty} c_k (1 - P_k(\cos\theta)) \tag{4.2}$$

where $P_k(\cdot)$ are Legendre polynomials, $\forall c_k \geq 0$ and $\sum c_k < \infty$.

It has been known that in $\mathbb{R}^d$, one can always obtain the variogram from the stationary covariance function with $\gamma(\theta) = C(0) - C(\theta)$ but not the converse. However, in S^2 [Yag61] argued that for a valid $\gamma(\theta), \theta \in [0, \pi]$ one can always construct the covariance function $C(\theta) = c_0 - \gamma(\theta)$ for some $c_0 \geq \int_0^\pi \gamma(\theta)\sin(\theta)d\theta$.

Here is the outline of this chapter. We first introduce axially symmetric random processes on the sphere and the representation for its covariance function. Next, we propose parametric models to generalize some of existing parametric models to capture the variation across latitudes when modeling the covariance structure of axially symmetric processes on the sphere. Finally, we discuss the properties of the cross-covariance and cross-variogram estimators based on Method of Moments.

4.2 Axial symmetry

For an axially symmetric process $X(P), P \in S^2$ on the sphere, the covariance function $R(P, Q)$ at two locations $P = (\phi_P, \lambda_P), Q(\phi_Q, \lambda_Q) \in S^2$ is given by

$$R(\phi_P, \phi_Q, \lambda_P, \lambda_Q) = R(\phi_P, \phi_Q, \lambda_P - \lambda_Q).$$

Following the discussion given in ([HZR12]), the covariance function can be expressed as the following summations.

$$\begin{aligned} R(P,Q) &= R(\phi_P, \phi_Q, \lambda_P - \lambda_Q) \\ &= \sum_{m=-\infty}^{\infty} \sum_{\nu=|m|}^{\infty} \sum_{\mu=|m|}^{\infty} f_{\nu,\mu,m} e^{im(\lambda_P - \lambda_Q)} P_\nu^m(\cos\phi_P) P_\mu^m(\cos\phi_Q), \end{aligned} \tag{4.3}$$

where $f_{\nu,\mu,m} = \overline{f}_{\mu,\nu,m}$ and for each fixed integer m, the matrix $F_m(N) = \{f_{\nu,\mu,m}\}_{\nu,\mu=|m|,|m|+1,\ldots,N}$ must be positive definite for all $N \geq |m|$. Furthermore, if we denote

$$C_m(\phi_P, \phi_Q) = \sum_{\nu=|m|}^{\infty} \sum_{\mu=|m|}^{\infty} f_{\nu,\mu,m} P_\nu^m(\cos\phi_P) P_\mu^m(\cos\phi_Q),$$

then

$$R(P,Q) = R(\phi_P, \phi_Q, \Delta\lambda) = \sum_{m=-\infty}^{\infty} e^{im\Delta\lambda} C_m(\phi_P, \phi_Q) \quad m = 0, \pm1, \pm2, \ldots \tag{4.4}$$

where $\Delta\lambda \in [-\pi, \pi]$ and $\phi_P, \phi_Q \in [0, \pi]$. Here the complex bivariate continuous function $C_m(\phi_P, \phi_Q)$ satisfies the following conditions.

- Hermitian and positive definite.
- $\sum_{m=-\infty}^{\infty} |C_m(\phi_P, \phi_Q)| < \infty$ for $m = 0, \pm1, \pm2, \ldots$ and any $\phi_P, \phi_Q \in [0, \pi]$.

One can use the inverse Fourier transformation to derive $C_m(\phi_P, \phi_Q)$ based on $R(P,Q)$,

$$C_m(\phi_P, \phi_Q) = \frac{1}{2\pi} \int_{-\pi}^{\pi} R(\phi_P, \phi_Q) e^{-im\Delta\lambda} d\Delta\lambda$$

Let $C_m(\phi_P, \phi_Q) = C_m^R(\phi_P, \phi_Q) + iC_m^I(\phi_P, \phi_Q)$. From [HZR12], if a random process is real-valued, its corresponding covariance function $R(P,Q)$ is also real-valued and so we have $C_m(\phi_P, \phi_Q) = \overline{C_{-m}(\phi_P, \phi_Q)}$. The covariance function $R(P,Q)$ on the sphere given by 4.4 can then be simplified as the following form.

$$\begin{aligned} R(P,Q) &= C_0(\phi_P,\phi_Q) + \sum_{m=1}^{\infty} e^{-im\Delta\lambda} C_{-m}(\phi_P,\phi_Q) + \sum_{m=1}^{\infty} e^{im\Delta\lambda} C_m(\phi_P,\phi_Q) \\ &= c_o C_0^R(\phi_P,\phi_Q) + 2\sum_{m=1}^{\infty} c_m[\cos(m\Delta\lambda) C_m^R(\phi_P,\phi_Q) - \sin(m\Delta\lambda) C_m^I(\phi_P,\phi_Q)]. \end{aligned}$$

Longitudinally reversible processes

If we further assume that the covariance function $R(P,Q)$ of an axially symmetric process satisfies the following.

$$R(\phi_P,\phi_Q,\lambda_P-\lambda_Q) = R(\phi_P,\phi_Q,\lambda_P-\lambda_Q), \tag{4.5}$$

we then call the underline process to be longitudinally reversible ([Ste07]). Now the covariance function $R(P,Q)$ of a real-valued longitudinally reversible process reduces to the following.

$$R(P,Q) = \sum_{m=0}^{\infty} C_m(\phi_P,\phi_Q)\cos(m\Delta\lambda)$$

as $C_m(\phi_P,\phi_Q)$ is real so that we have $C_{-m}(\phi_P,\phi_Q) = C_m(\phi_P,\phi_Q)$ and $C_m(\phi_P,\phi_Q) = \overline{C_{-m}(\phi_P,\phi_Q)}$.

4.3 Proposed parametric models

As discussed in Chapter 2, many covariance models that are valid in R^d might not be valid in S^2. Therefore, it is necessary to develop parametric models that capture the topological structure of compactness of a sphere. Although the Matérn covariance model has been often used on modeling the global data in recent years, it has some serious drawbacks. For example, it has been shown that the homogeneous Matérn covariance model is not valid when the smoothness parameter is bigger than 0.5. Further modifications have been proposed in various researches, see [Li13], [JS08], [?], to name a few. [HZR12] discussed a new representation for the covariance structure of an axially symmetric process, and based on the parametric form of $C_m(\phi_P, \phi_Q)$, they proposed some parametric covariance models.

The covariance function on sphere, $R(P, Q)$, given in equation 4.4, is clearly a function of both longitude and latitude.

$$R(P, Q) = f(\Delta\lambda, \phi_P, \phi_Q)$$

In order to make things easier one could assume that $C_m(\phi_P, \phi_Q) = \tilde{C}_m(\phi_P - \phi_Q)$ only depends on the difference of ϕ_P and ϕ_Q, [HZR11] proposed a simple separable covariance function when both covariance components are exponential

$$R(P, Q) = c_0 e^{-a|\Delta\lambda|} e^{-b|\phi_P - \phi_Q|},$$

Where a and b are defined as decay parameters in longitude and latitude respectively. The separable models are too simple and they are not capable to capture the covariance structure of the entire sphere. Therefore, [HZR12] proposed some non-separable covariance models by carefully choosing functions for $C_m(\phi_P, \phi_Q)$ that are valid on the sphere,

$$R(P,Q) = Ce^{-a|\phi_P-\phi_Q|}\frac{1-p^2}{1-2p\cos\Theta+p^2} \tag{4.6}$$

$$R(P,Q) = Ce^{-a|\phi_P-\phi_Q|}\log\frac{1}{(1-2p\cos\Theta*+p^2)} \tag{4.7}$$

$$R(P,Q) = 2Ce^{-a|\phi_P-\phi_Q|}\left(\frac{\pi^4}{90}-\frac{\pi^2\Theta^2}{12}+\frac{\pi\Theta^3}{12}-\frac{\Theta^4}{48}\right), \tag{4.8}$$

where $\Theta = \Delta\lambda + u(\phi_P - \phi_Q) - 2k\pi$, and k is chosen such that $\Theta \in [0, 2\pi]$.

There is one big disadvantage of the covariance models proposed by [HZR12], the biggest disadvantage for all of them are that it is assumed not only stationarity on longitudes, but stationarity on latitudes as well.

We have noticed that when $\phi_P = \phi_Q$, the model 4.6 reduces to

$$R(P,P) = C\frac{1-p^2}{1-2p\cos(\Delta\lambda)+p^2}$$

and if we set $\Delta\lambda = 0$, the variance of latitude ϕ_P over all latitudes can be given by,

$$Var(P) = C\frac{1+p}{1-p}$$

is not a function of the latitude (a function of the parameter p) and it implies that variance is constant over all latitudes. This is not supposed to be the case, since both MSU data and TOMS data in figures 2b and 4b shows that variance is highly depending on the latitude. In order to overcome this issue we propose a solution to the above covariance models to capture the non stationarity over the latitudes.

Proposition 4.1. *A more general non stationary covariance function is given as following. If $C(\cdot) = C(x-y)$ is the stationary covariance function and $f(\omega) \geq 0$ is the corresponding spectral density, then*

$$\tilde{C}(x,y) = C_2 - C(x) - C(y) + C(x-y),$$

with

$$C_2 \geq \int_{-\infty}^{\infty} dF(\omega) = \int_{-\infty}^{\infty} f(\omega) d\omega > 0$$

is the non stationary covariance function. Note that the covariance function $C(\cdot)$ implies that, by Bochner's theorem, there exists a bounded measure F such that

$$C(x) = \int_{-\infty}^{\infty} e^{-ix\omega} dF(\omega).$$

When $F(\cdot)$ is absolutely continuous, there exists a spectral density $f(\cdot) \geq 0$ such that

$$C(x) = \int_{-\infty}^{\infty} e^{-ix\omega} f(\omega) d\omega.$$

Now we choose a sequence of complex numbers $a_i, i = 1, 2, \cdots, n$, *and any sequence of real numbers* $t_i, i = 1, 2, \cdots, n$, *taking* $C_2 = \int_{-\infty}^{\infty} f(\omega)d\omega$,

$$\begin{aligned}
\sum_{i=1}^{n}\sum_{j=1}^{n} a_i\overline{a}_j\tilde{C}(t_i, t_j) &= \sum_{i}\sum_{j} a_i\overline{a}_j(C_2 - C(t_i) - C(-t_j) + C(t_i - t_j)) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i\overline{a}_j \int_{-\infty}^{\infty} (1 - e^{-it_i\omega} - e^{it_j\omega} + e^{-i(t_i-t_j)\omega})f(\omega)d\omega \\
&= \int_{-\infty}^{\infty} f(\omega)d\omega \left|\sum_{i=1}^{n} a_i(e^{-it_i\omega} - 1)\right|^2 \geq 0.
\end{aligned}$$

In the case of circle clearly $C_1e^{|\theta|}$ is a stationary covariance function and we can apply 4.1 to get a non stationary covariance function. Lets consider the following stationary covariance functions over the latitudes,

$$C(\phi) = Ce^{-a|\phi_P|}$$
$$C(\phi) = C\frac{1}{\sqrt{a^2 + \phi^2}}$$

Now, we apply proposition 4.1 to get non-stationary covariance functions, which depends on the latitudes, even when $\phi_P = \phi_Q$. Consider the below two functions for $C_m(\phi_P, \phi_Q)$.

$$\tilde{C}(\phi_P, \phi_Q) = C_1(C_2 - e^{-a|\phi_P|} - e^{-a|\phi_Q|} + e^{-a|\phi_P-\phi_Q|}) \tag{4.9}$$
$$\tilde{C}(\phi_P, \phi_Q) = C_1\left(C_2 - \frac{1}{\sqrt{a^2 + \phi_P^2}} - \frac{1}{\sqrt{a^2 + \phi_Q^2}} + \frac{1}{\sqrt{a^2 + (\phi_P - \phi_Q)^2}}\right) \tag{4.10}$$

Here $C_1, a > 0$, and $C_2 \geq 1$ to ensure the positive definiteness of the above function. When $\phi_P = \phi_Q$, both functions are actually a function of ϕ_P.

$$\begin{aligned}
\tilde{C}(\phi_P, \phi_P) &= C_1(C_2 - 2e^{-a|\phi_P|} + 1), \\
\tilde{C}(\phi_P, \phi_P) &= C_1\left(C_2 - \frac{2}{\sqrt{a^2 + \phi_P^2}} + \frac{1}{a}\right).
\end{aligned}$$

So we propose six five-parameter models which are combinations of both $\tilde{C}(\phi_P, \phi_Q)$, defined by a exponential family 4.9 and a power family 4.10, and models (4.6, 4.7, 4.8) proposed by [HZR12] for the covariance on a sphere defined as follows,

$$R(P, Q) = \tilde{C}(\phi_P, \phi_Q)C(\Theta),$$

where $\Theta = \Delta\lambda + u(\phi_P - \phi_Q) \in [0, 2\pi]$, $C_1 > 0, C_2 > 0, a > 0, u \in \mathbb{R}, p \in (0, 1)$.

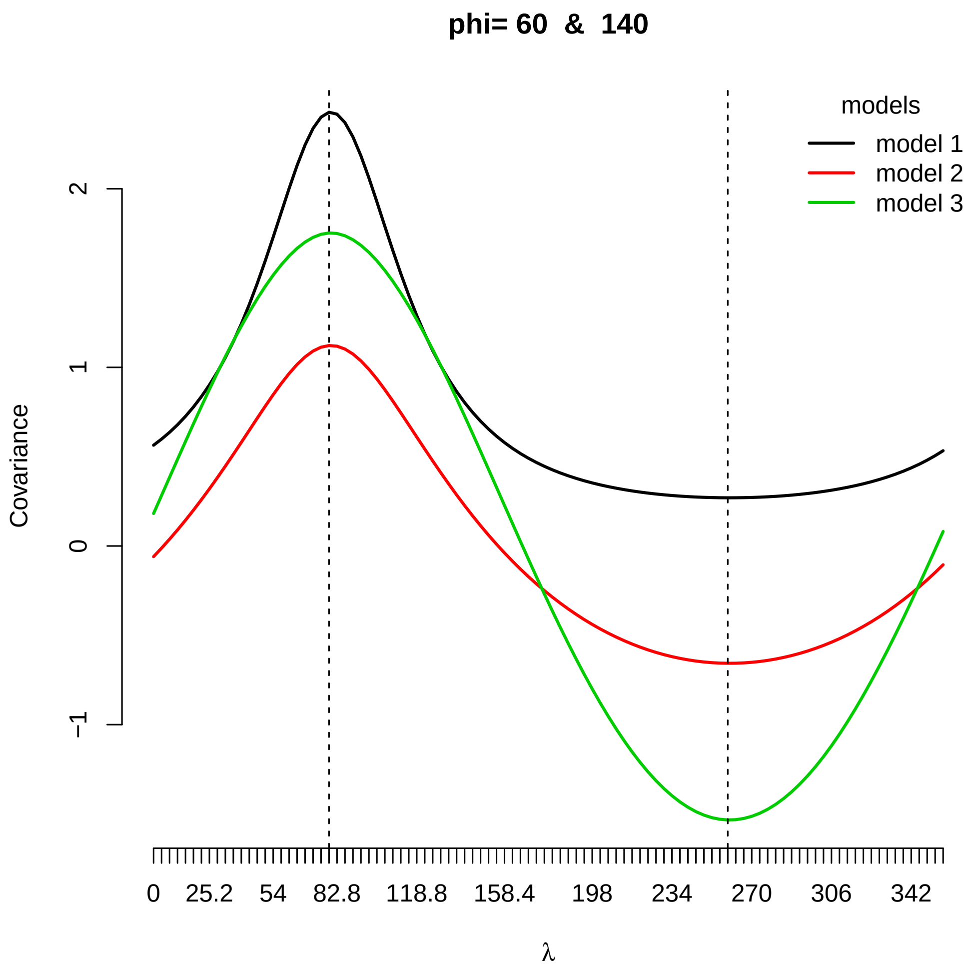


Figure 10. The covariance between $30^0 S$ and $50^0 N$ (latitude 60^0 and 140^0) of three covariance models with exponential family $i.e.\tilde{C}(\phi_P, \phi_Q)$ given by 4.9 over 100 longitudes for simplicity we set all parameters to be one.

Remark 1 The parameters C_1, C_2, a, p are scaling parameters of the covariance functions and u is a location parameter. All covariance models have a similar pattern and share one property, when there is no location shift $(u = 1)$ the maximum of

$R(P,Q)$ occurs at $\lambda_{max} = |\phi_P - \phi_Q|$ and the minimum of $R(P,Q)$ occurs at $\lambda_{min} = \pi + \lambda_{max}$.

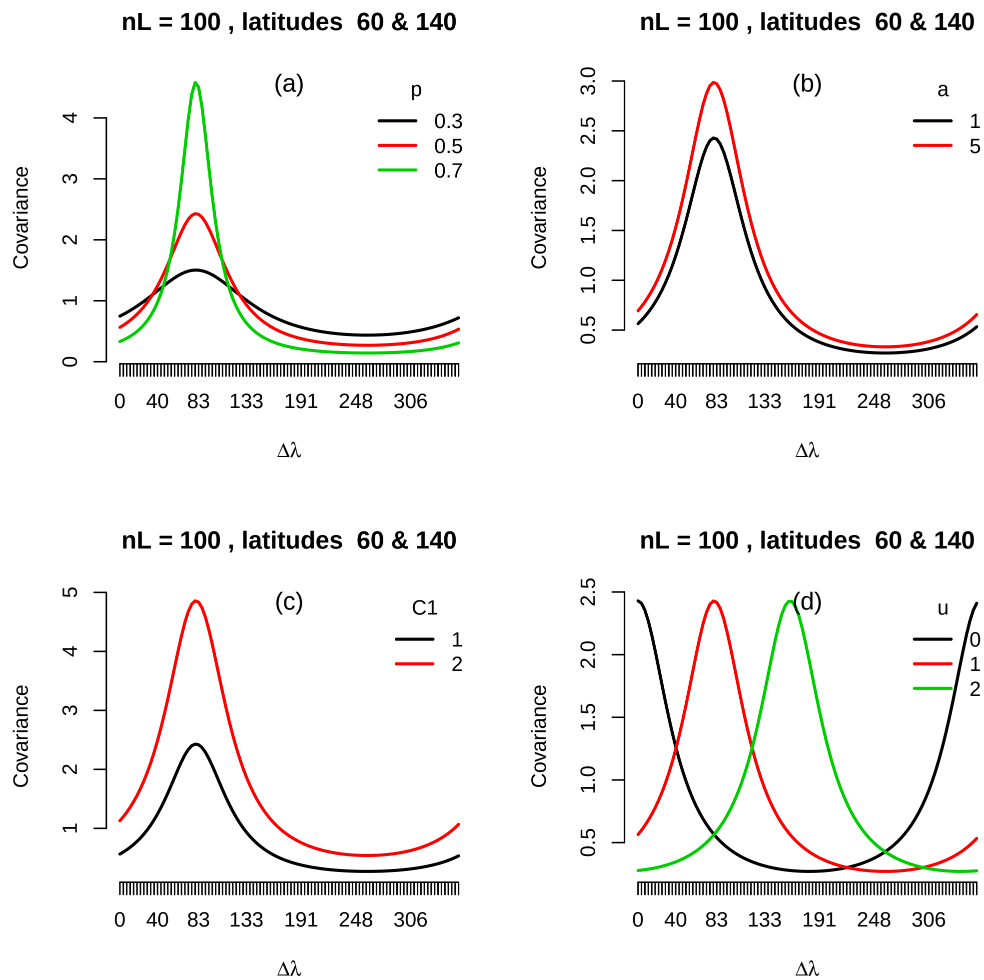


Figure 11. Covariance distribution for different parameter using model1: (a)-parameter p, (b)-parameter a, (c)-parameter $C1$ (similar pattern for parameter $C2$), (d)-parameter u

Remark 2 The scaling parameter p is more sensitive at supremum and infimum of the covariance models and parameters C_1, C_2, a are regular scaling parameters.

The parameter u is a location parameter which shits the covariance from left to right ($\Delta\lambda$) when $u > 0$, but $u = 0$ will provide a longitudinally reversible covariance model which is similar to Matérn covariance model when the smoothing parameter (ν) is $1/2$.

4.4 Covariance and variogram estimators on a sphere

Cross Covariance

The cross covariance captures the covarince between two locations and any finite pairs of locations separated by a fixed distance (longitudinal difference $\Delta\lambda$). In other words cross covariance can be used to to capture the covariance between points at two latitudes separated by $\Delta\lambda \in (0, 2\pi)$. When a axially symmetric random process on a sphere is second-order stationary, cross covariance is a function of longitudinal difference ($\Delta\lambda$). According to [Wac13] the cross covariance function is not an even function and it is easy to observe that the proposed $R(P, Q)$ functions are valid on a sphere and they are cross covarince functions ($R(P, Q, \Delta\lambda) \neq R(P, Q, -\Delta\lambda)$). The cross covariance estimate for axially symmetric processes on the sphere. For any two latitudes ϕ_P and ϕ_Q with $\{\lambda_i, i = 1, 2, \ldots, n\}$ representing the gridded longitudes on each circle, then $\hat{R}(\phi_P, \phi_Q, \Delta\lambda)$ is given by

$$\hat{R}(\phi_P, \phi_Q, \Delta\lambda) = \frac{1}{n}\sum_{i=1}^{n}(X(\phi_P, \lambda_i + \Delta\lambda) - \bar{X}_P)(X(\phi_Q, \lambda_i) - \bar{X}_Q), \tag{4.11}$$

where $\Delta\lambda = 0, 2\pi/n, 4\pi/n, \cdots, 2(N-1)\pi/n$ and $\bar{X}_P = \frac{1}{n}\sum_{i=1}^{n} X(\phi_P, \lambda_i)$ and similar for $\bar{X}_Q$. Now we calculate the unbiasedness of the cross covarince estimator.

$$\begin{aligned}
E(\hat{R}(\phi_P, \phi_Q, \Delta\lambda)) &= \frac{1}{n}\sum_{i=1}^{n} E((X(\phi_P, \lambda_i + \Delta\lambda) - \bar{X}_P)(X(\phi_Q, \lambda_i) - \bar{X}_Q)) \\
&= \frac{1}{n}\sum_{i=1}^{n} cov(X(\phi_P, \lambda_i + \Delta\lambda), X(\phi_Q, \lambda_i)) \\
&\quad -\frac{1}{n}\sum_{i=1}^{n} E((X(\phi_P, \lambda_i + \Delta\lambda) - \mu_P)(\bar{X}_Q - \mu_Q)) \\
&\quad -\frac{1}{n}\sum_{i=1}^{n} E((X(\phi_Q, \lambda_i) - \mu_Q)(\bar{X}_P - \mu_P)) \\
&\quad +\frac{1}{n}\sum_{i=1}^{n} E((\bar{X}_P - \mu_P)(\bar{X}_Q - \mu_Q)) \\
&= R(\phi_P, \phi_Q, \Delta\lambda) - E((\bar{X}_Q - \mu_Q)(\bar{X}_P - \mu_P)) - E((\bar{X}_P - \mu_P)(\bar{X}_Q - \mu_Q)) \\
&\quad +E((\bar{X}_P - \mu_P)(\bar{X}_Q - \mu_Q)) \\
&= R(\phi_P, \phi_Q, \Delta\lambda) - cov(\bar{X}_P, \bar{X}_Q).
\end{aligned}$$

Note that,

$$
\begin{aligned}
cov(\bar{X}_P, \bar{X}_Q) &= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} cov(X(\phi_P,\lambda_i), X(\phi_Q,\lambda_j)) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} R(\phi_P,\phi_Q,(i-j)*2\pi/n) \\
&= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} (C_0(\phi_P,\phi_Q) \\
& \quad 2\sum_{m=1}^{\infty}(C_{m,R}(\phi_P,\phi_Q)\cos(m*(i-j)*2\pi/n) \\
& \quad -C_{m,I}(\phi_P,\phi_Q)\sin(m*(i-j)*2\pi/n)) \\
&= C_0(\phi_P,\phi_Q) + 2\sum_{m=1}^{\infty} C_{m,R}(\phi_P,\phi_Q)\left(\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\cos(m(i-j)*2\pi/n)\right) \\
& \quad -2\sum_{m=1}^{\infty} C_{m,I}(\phi_P,\phi_Q)\left(\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sin(m(i-j)*2\pi/n)\right) \\
&= C_0(\phi_P,\phi_Q),
\end{aligned}
$$

since

$$
\begin{aligned}
& \sum_{i=1}^{n}\sum_{j=1}^{n}\cos(m*(i-j)*2\pi/n) \\
=\ & \sum_{i=1}^{n}\sum_{j=1}^{n}(\cos(m*i*2\pi/n)\cos(m*j*2\pi/n) - \sin(m*i*2\pi/n)\sin(m*j*2\pi/n)) \\
=\ & \left(\sum_{i=1}^{n}\cos(m*i*2\pi/n)\right)^2 - \left(\sum_{i=1}^{n}\sin(m*i*2\pi/n)\right)^2 = 0
\end{aligned}
$$

and

$$
\begin{aligned}
& \sum_{i=1}^{n}\sum_{j=1}^{n}\sin(m*(i-j)*2\pi/n) \\
= & \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sin(m*i*2\pi/n)\cos(m*j*2\pi/n)-\cos(m*i*2\pi/n)\sin(m*j*2\pi/n)\right) \\
= & \left(\sum_{i=1}^{n}\cos(m*i*2\pi/n)\right)*\left(\sum_{i=1}^{n}\sin(m*i*2\pi/n)\right) \\
& -\left(\sum_{i=1}^{n}\cos(m*i*2\pi/n)\right)*\left(\sum_{i=1}^{n}\sin(m*i*2\pi/n)\right)=0
\end{aligned}
$$

since for any integer m, we have

$$
\sum_{k=1}^{n}\cos(mk*2\pi/n)=\begin{cases}0, & \text{for any integer } m\neq 0,\\ n, & \text{for } m=0\end{cases} \quad \text{and} \quad \sum_{k=1}^{n}\sin(mk*2\pi/n)=0.
$$

Hence,

$$
cov(\bar{X}_P,\bar{X}_Q)=C_0(\phi_P,\phi_Q).
$$

Therefore,

$$
E(\hat{R}(\phi_P,\phi_Q,\Delta\lambda))=R(\phi_P,\phi_Q,\Delta\lambda)-C_0(\phi_P,\phi_Q).
$$

The cross covariance estimator is biased and when $\phi_P=\phi_Q$ this reduces to the same results we obtained for a random process on the circle.

Remark 3 If the mean at each latitude is zero the above cross covariance estimator can be given by

$$\hat{R}(\phi_P, \phi_Q, \Delta\lambda) = \frac{1}{n}\sum_{i=1}^{n} X(\phi_P, \lambda_i + \Delta\lambda) X(\phi_Q, \lambda_i)$$

is unbiased. However, this estimator may not be practically reasonable in axially symmetric geo-spatial systems as the mean is a function of latitude.

Cross variogram

In general for a stationary process when the covariance is known one can get the variogram ($2\gamma(\theta) = C(0) - C(\theta)$), since the cross covariance is not an even function the variogram is defined by taking the average of $R(P, Q, \Delta\lambda)$ and $R(P, Q, -\Delta\lambda)$ and we can derive the cross variogram as follows,

$$\begin{aligned}
\gamma(\phi_p, \phi_Q, \Delta\lambda) &= \frac{1}{2}E\left((X(\phi_P, \lambda + \Delta\lambda) - X(\phi_P, \lambda))(X(\phi_Q, \lambda + \Delta\lambda) - X(\phi_Q, \lambda))\right) \\
&= \frac{1}{2}E\left(((X(\phi_P, \lambda + \Delta\lambda) - \mu_P) - (X(\phi_P, \lambda) - \mu_P))\right. \\
&\quad \left.((X(\phi_Q, \lambda + \Delta\lambda) - \mu_Q) - (X(\phi_Q, \lambda) - \mu_Q)\right) \\
&= \frac{1}{2}\left(cov(X(\phi_P, \lambda + \Delta\lambda), X(\phi_Q, \lambda + \Delta\lambda)) - cov(X(\phi_P, \lambda + \Delta\lambda), X(\phi_Q, \lambda))\right. \\
&\quad \left. - cov(X(\phi_P, \lambda), X(\phi_Q, \lambda + \Delta\lambda)) + cov(X(\phi_P, \lambda), X(\phi_Q, \lambda))\right) \\
&= \frac{1}{2}\left(R(\phi_P, \phi_Q, 0) - R(\phi_P, \phi_Q, \Delta\lambda) - R(\phi_P, \phi_Q, -\Delta\lambda) + R(\phi_P, \phi_Q, 0)\right) \\
&= R(\phi_P, \phi_Q, 0) - \frac{1}{2}(R(\phi_P, \phi_Q, \Delta\lambda) + R(\phi_P, \phi_Q, -\Delta\lambda)).
\end{aligned}$$

$$\gamma(\phi_p, \phi_Q, \Delta\lambda) = R(\phi_p, \phi_Q, 0) - \frac{1}{2}(R(\phi_P, \phi_Q, \Delta\lambda) + R(\phi_P, \phi_Q, -\Delta\lambda)). \tag{4.12}$$

The MOM estimator for cross-variogram on axially symmetric processes on the sphere is given by

$$\hat{\gamma}(\phi_p, \phi_Q, \Delta\lambda) = \frac{1}{2n}\sum_{i=1}^{n}\left(X(\phi_P, \Lambda_i) - X(\phi_P, \lambda_i))(X(\phi_Q, \Lambda_i) - X(\phi_Q, \lambda_i))\right), \tag{4.13}$$

where $\Lambda_i = \lambda_i + \Delta\lambda$ and we have

$$\begin{aligned} E(\hat{\gamma}_{PQ}(\Delta\lambda)) &= \frac{1}{2n}\sum_{i=1}^{n} E\left(X(\phi_P, \lambda_i + \Delta\lambda) - X(\phi_P, \lambda_i))(X(\phi_Q, \lambda_i + \Delta\lambda) - X(\phi_Q, \lambda_i))\right) \\ &= \frac{1}{2n}\sum_{i=1}^{n}\left(2\gamma(\phi_p, \phi_Q, \Delta\lambda)\right) = \gamma(\phi_p, \phi_Q, \Delta\lambda), \end{aligned}$$

which is unbiased.

Remark 4 Lets re arrange $R(P, Q)$,

$$R(\phi_P, \phi_Q, \Delta\lambda) = \frac{1}{2}(R(\phi_P, \phi_Q, \Delta\lambda) + R(\phi_P, \phi_Q, -\Delta\lambda)) + \frac{1}{2}(R(\phi_P, \phi_Q, \Delta\lambda) - R(\phi_P, \phi_Q, -\Delta\lambda))$$

the cross-covariance function $R(\phi_P, \phi_Q, \Delta\lambda)$ is decomposed into two components: the even component (the first average) and the odd component (the second average).

The cross-variogram is only related to the even component of the cross-covariance function, which is different from the case on the circle (the covariance is an even function). [Wac13] argues that cross variogram is not sufficient when there is a delayed affect. However, in the data generation process there is no delayed affect (between latitudes).

CHAPTER V

GLOBAL DATA GENERATION AND ESTIMATION ON THE SPHERE

In Chapter 4 we discussed about random process on the sphere and so far very few studies have attempted to model global data on a unit sphere ($\mathbb{S}^2$). The previous studies have argued that many processes on a sphere are not homogeneous, especially in the direction of latitude ([Ste07]). In order to capture non stationarity [JS07] proposed spatio-tempo covariance functions on the sphere by applying first order differential operator to fully symmetric spatio-tempo processes on sphere. Further, [JS08] proposed flexible class of parametric covariance models to capture the non-stationarity of global data. Discrete Fourier Transform (DFT) was used for the data on regular grids and calculated the exact likelihood for large data sets. Furthermore, they used Legendre polynomials to remove the spatial trends when fitting models to global data ([Ste07]). Li [Li13] discussed about the issues associated when modeling axially symmetric spatial random fields on a sphere. Further, They proposed convolution methods to generate random fields with a class of $Mat\acute{e}rn$-type kernel functions by allowing the parameters in the kernel function to vary with latitude. Moreover, they were able to generate flexible class of covariance functions and capture the non-stationary properties on a sphere. It is a well known phenomenon that inverting a $n \times n$ matrix cost $\mathcal{O}(n^3)$. However in spatial spatial systems the covariance matrices are very large and inverting these matrices could be beyond the current computational resources. For example the MSU data was observed on a $72^0 \times 144^0$ grid which result in a covarince matrix with a dimension 10368×10368. One could use the properties of covariance matrix, block circulant, to compute the inverse of covarince matrix ([DMG83, ?, Tee05, JS08, Li13]). Our proposed covarince models $R(P,Q)$ are in fact block circulant matrices and the inverting cost is $\mathcal{O}(n_L {n_l}^3)$ ([Li13]) where n_L is the number of longitudes and n_l is the number of latitudes.

The global data generation process on a sphere discussed on this dissertation is primarily based on the axially symmetric covariance structure introduced by [Jon63] and as continuation of axially symmetric process on a sphere developed by [HZR12]. In this chapter, we first layout the details and methodologies of generating data on a sphere using a circularly-symmetric matrices. Then we provide a pseudo code for global data generation process and discuss about the simulation setup. Finally, we use cross variogram and cross covariance to compare the simulated data.

Let $X(P)$ be a complex-valued random process defined on a unit sphere S^2, where $P = (\lambda, \phi) \in S^2$ with longitude $\lambda \in [-\pi, \pi)$ and latitude $\phi \in [0, \pi]$. In chapter 4 we discussed how to formulate a valid covariance function for continuous axially symmetric processes on a sphere and was given by 4.4. Now, in the light of [HZR12][remark 2.5] a continuous axially symmetric process, $X(P)$ on a unit sphere , is given by

$$X(P) = X(\phi, \lambda) = \sum_{m=-\infty}^{\infty} W_{m,\nu}(\phi) e^{im\lambda} \psi_{m,\nu}(\phi), \tag{5.1}$$

where λ is the longitude, ϕ is the latitude and $\psi_{m,\nu}(\cdot)$ is a orthonormal basis for $C_m(\phi_P, \phi_Q)$ and using inverse Fourier transformation we get

$$W_m(\phi) = \frac{1}{2\pi} \int_{S^2} X(P) e^{-im\lambda} \overline{\psi_{m,\nu}(\phi)} dP,$$

with $\mathrm{E}(W_{m,\nu}\overline{W_{n,\nu}}) = \delta_{m,n}\delta_{m,\mu}\eta_{m,\nu}$.

Remark 1 If the random process on a sphere is real-Gaussian, then the weights given by $W_{m,\nu}$ will be independent normal random variables. Furthermore, if ν is fixed the process defined by $X(P)$ will be equivalent to a homogeneous random process on a circle with angular distance $\Delta\lambda$. In other words a random process on a sphere at a given latitude (for fix ϕ) can be studies as a random process on the circle. In Chapter 3 a random process on a circle was given by a infinite Fourier summation ([Roy72],[DR76]) and in the case of a process on a circle, $X(P)$ (5.1) can be given by

$$X(\phi, \lambda) = \sum_{m=-\infty}^{\infty} W_m(\phi) e^{im\lambda}, \tag{5.2}$$

$$\text{where } W_m(\phi) = \frac{1}{2\pi} \int_0^{2\pi} X(\phi, \lambda) e^{-im\lambda} d\lambda,$$

with $\mathrm{E}(W_m(\phi_P)\overline{W_n(\phi_Q)}) = \delta_{m,n} C_m(\phi_P, \phi_Q)$.

5.1 Method development

We can construct normal independent (complex) random variate $W_m(\phi)$ associated with the variance-covariance matrix $C_m(\phi_P, \phi_Q)$ to construct an axially symmetric process for a given latitude ϕ. Then finite summation can be used to approximate above (5.2) infinite summation as given below,

$$X(P) = X(\phi, \lambda) = \sum_{m=-N}^{N} W_m(\phi) e^{im\lambda} \tag{5.3}$$

where this would provide the gridded data. Since W_m's are independent for $m = 1, 2, \cdots$, we have

$$\begin{aligned} Cov(X(P), X(Q)) &= Cov\left(\sum_{m=-N}^{N} W_m(\phi_P) e^{im\lambda_P}, \sum_{j=-N}^{N} W_j(\phi_Q) e^{ij\lambda_Q}\right) \\ &= \sum_{m,j} e^{im\lambda_P} e^{-ij\lambda_Q} Cov(W_m(\phi_P), W_j(\phi_Q)) \\ &= \sum_{m} e^{im(\lambda_P - \lambda_Q)} C_m(\phi_P, \phi_Q) \end{aligned}$$

The above generated data will be complex random variates. Therefore to have the real-valued data observations or to obtain a real process, we need to have

$$C_{-m}(\phi_P, \phi_Q) = \overline{C_m(\phi_P, \phi_Q)}, \quad \text{for } m = 1, 2, \cdots, N \tag{5.4}$$

Lets write $W_m(\phi) = W_m^r(\phi) + iW_m^i(\phi)$ in terms of a real component and an imaginary component. We also write $C_m(\phi_P, \phi_Q) = C_m^r(\phi_P, \phi_Q) + iC_m^i(\phi_P, \phi_Q)$ and with the relationship 5.4 above, we have

$$C_{-m}^r(\phi_P, \phi_Q) = C_{-m}^r(\phi_P, \phi_Q), \quad C_{-m}^i(\phi_P, \phi_Q) = -C_{-m}^i(\phi_P, \phi_Q).$$

Now,

$$\begin{aligned}
Cov(W_m(\phi_P), W_m(\phi_Q)) &= Cov(W_m^r(\phi_P) + iW_m^i(\phi_P), W_m^r(\phi_Q) + iW_m^i(\phi_Q)) \\
&= \left[Cov(W_m^r(\phi_P), W_m^r(\phi_Q)) + Cov(W_m^i(\phi_P), W_m^i(\phi_Q))\right] \\
&\quad +i\left[-Cov(W_m^r(\phi_P), W_m^i(\phi_Q)) + Cov(W_m^i(\phi_P), W_m^r(\phi_Q))\right] \\
&= C_m^r(\phi_P, \phi_Q) + iC_m^i(\phi_P, \phi_Q).
\end{aligned}$$

If we let $W_{-m}(\phi) = \overline{W_m(\phi)}$, then the covariance function would satisfy the above relationship 5.4. In addition, we will set the following,

$$Cov(W_m^r(\phi_P), W_m^r(\phi_Q)) = Cov(W_m^i(\phi_P), W_m^i(\phi_Q)) = \frac{1}{2}C_m^r(\phi_P, \phi_Q), \quad (5.5)$$

$$Cov(W_m^i(\phi_P), W_m^r(\phi_Q)) = -Cov(W_m^r(\phi_P), W_m^i(\phi_Q)) = \frac{1}{2}C_m^i(\phi_P, \phi_Q). \quad (5.6)$$

Therefore, if we denote $\underset{\sim}{W}_m(\phi) = (W_m^r(\phi), W_m^i(\phi))^T$, then the variance-covariance matrix for $\underset{\sim}{W}_m(\phi)$ is given by

$$\frac{1}{2}\begin{pmatrix} C_m^r(\phi_P, \phi_Q) & -C_m^i(\phi_P, \phi_Q) \\ C_m^i(\phi_P, \phi_Q) & C_m^r(\phi_P, \phi_Q) \end{pmatrix}.$$

However, we cannot have a vector of random variables $\underset{\sim}{W}_m(\phi)$ with a non-symmetric variance-covariance matrix unless $C_m^i(\phi_P, \phi_Q) = 0$. In the next section we will demonstrate how to generate $\underset{\sim}{W}_m(\phi)$ with a symmetric variance-covariance

The process given by (5.2) is now simplified as the following (real) process,

$$\begin{aligned} X(P) &= \sum_{m=-N}^{N} W_m(\phi)e^{im\lambda} = W_0(\phi) + \sum_{m=1}^{N} W_m(\phi)e^{im\lambda} + \sum_{m=-1}^{-N} W_m(\phi)e^{im\lambda} \\ &= W_0(\phi) + \sum_{m=1}^{N} W_m(\phi)e^{im\lambda} + \sum_{m=1}^{N} \overline{W_m(\phi)}e^{-im\lambda} \\ &= W_0(\phi) + \sum_{m=1}^{N} \big[(W_m^r(\phi) + iW_m^i(\phi))(\cos(m\lambda) + i\sin(m\lambda)) \\ &\quad + (W_m^r(\phi) - iW_m^i(\phi))(\cos(m\lambda) - i\sin(m\lambda))\big] \\ &= W_0(\phi) + 2\sum_{m=1}^{N} \big[W_m^r(\phi)\cos(m\lambda) - W_m^i(\phi)\sin(m\lambda)\big]. \end{aligned} \tag{5.7}$$

Data generation

Now for each fixed $m = 0, 1, 2, \cdots, N$, we consider $W_m(\phi) = W_m^r(\phi) + iW_m^i(\phi)$ then $W_m^*(\phi) = W_m^r(\phi) - iW_m^i(\phi)$ (where $W_m^*(\phi)$ is the complex conjugate of $W_m(\phi)$). We may assume that $W_m^r(\phi)$ and $W_m^i(\phi)$ are independent, each following a (Gaussian) distribution with mean zero and the same variance $\sigma_m^2(\phi) = \frac{1}{2}C_m^r(\phi,\phi)$, ($C_m^i(\phi,\phi) = 0$ implies $W_m^r(\phi)$ and $W_m^i(\phi)$ are uncorrelated, or independent for Gaussian). In chapter 1 we introduced the concept of circularly-symmetry, thus according to [Gal08] a complex random variable is circularly-symmetric if and only if it pseudo covariance is zero (1.2). In this section we will show that the Gaussian random variable $W_m(\phi)$ is a circularly-symmetric complex random variable.

Now for a set of distinct latitudes $\Phi = \{\phi_1, \phi_2, \cdots, \phi_{n_l}\}$, we consider a sequence of complex random variables $\{W_m(\phi) : \phi \in \Phi\}$, which forms a multivariate complex random vector $\underset{\sim}{W}_m = (W_m(\phi_1), W_m(\phi_2), \cdots, W_m(\phi_n))^T$ where $W_m(\phi_i) = W_m^r(\phi_i) + iW_m^r(\phi_i)$ with associated $2 \times n_l$-dimensional real random vector

$$\underset{\sim}{V}_m = (W_m^r(\phi_1), W_m^i(\phi_1), W_m^r(\phi_2), W_m^i(\phi_2), \cdots, W_m^r(\phi_{n_l}), W_m^i(\phi_{n_l}))^T.$$

Now we calculate the covariance matrix $K_W = E(\underset{\sim}{W}_m \underset{\sim}{W}_m^*)$ (where $\underset{\sim}{W}_m^*$ is the conjugated transpose) and pseudo-covariance $M_W = E(\underset{\sim}{W}_m \underset{\sim}{W}_m^T)$. Further, from 1.2 a complex random vector is circularly-symmetric if and only if M_W is zero.

$$\begin{aligned} M_W &= \begin{pmatrix} E[W_m(\phi_1)W_m(\phi_1)] & E[W_m(\phi_1)W_m(\phi_2)] & \cdots & E[W_m(\phi_1)W_m(\phi_{n_l})] \\ E[W_m(\phi_2)W_m(\phi_1)] & E[W_m(\phi_2)W_m(\phi_2)] & \cdots & E[W_m(\phi_2)W_m(\phi_{n_l})] \\ \vdots & \vdots & \ddots & \vdots \\ E[W_m(\phi_{n_l})W_m(\phi_1)] & E[W_m(\phi_{n_l})W_m(\phi_2)] & \cdots & E[W_m(\phi_{n_l})W_m(\phi_{n_l})] \end{pmatrix} \\ &= \mathbf{0} \end{aligned}$$

We can show the above result for $\forall\ i, j$,

$$
\begin{aligned}
& E[W_m(\phi_i)W_m(\phi_j)] \\
= \;& E[(W_m^r(\phi_i) + iW_m^i(\phi_i))(W_m^r(\phi_j) + iW_m^i(\phi_j))] \\
= \;& E(W_m^r(\phi_i)W_m^r(\phi_j)) - E(W_m^i(\phi_i)W_m^i(\phi_j)) + i[E(W_m^r(\phi_i)W_m^i(\phi_j)) + E(W_m^i(\phi_i)W_m^r(\phi_j))] \\
& \text{for } i \neq j \\
= \;& \frac{1}{2}(C_m^r(\phi_i, \phi_j) - C_m^r(\phi_i, \phi_j)) + i[-\frac{1}{2}C_m^i(\phi_i, \phi_j) + \frac{1}{2}C_m^i(\phi_i, \phi_j)] = 0 \\
& \text{for } i = j \\
= \;& \frac{1}{2}(C_m^r(\phi_i, \phi_i) - C_m^r(\phi_i, \phi_i)) + i[0 + 0] = 0 \quad ; W_m^r(\phi_i), W_m^i(\phi_i) \text{ are independent}
\end{aligned}
$$

Therefore, $\underset{\sim}{W}_m$ is circularly-symmetric. In addition,

$$
\begin{aligned}
K_W &= E(\underset{\sim}{W}_m \underset{\sim}{W}_m^*) \\
&= \begin{pmatrix} E[W_m(\phi_1)W_m^*(\phi_1)] & E[W_m(\phi_1)W_m^*(\phi_2)] & \cdots & E[W_m(\phi_1)W_m^*(\phi_{n_l})] \\ E[W_m(\phi_2)W_m^*(\phi_1)] & E[W_m(\phi_2)W_m^*(\phi_2)] & \cdots & E[W_m(\phi_2)W_m^*(\phi_{n_l})] \\ \vdots & \vdots & \ddots & \vdots \\ E[W_m(\phi_{n_l})W_m^*(\phi_1)] & E[W_m(\phi_{n_l})W_m^*(\phi_2)] & \cdots & E[W_m(\phi_{n_l})W_m^*(\phi_{n_l})] \end{pmatrix} \\
&= \begin{pmatrix} C_{11}^r & C_{12}^r + iC_{12}^i & \cdots & C_{1n_l}^r + iC_{1n_l}^i \\ C_{21}^r - iC_{21}^i & C_{22}^r & \cdots & C_{2n_l}^r + iC_{2n_l}^i \\ \vdots & \vdots & \ddots & \vdots \\ C_{n_l1}^r - iC_{n_l1}^i & C_{n_l2}^r - iC_{n_l2}^i & \cdots & C_{n_ln_l}^r \end{pmatrix} \\
&= \begin{pmatrix} C_{11}^r & C_{12}^r & \cdots & C_{1n_l}^r \\ C_{21}^r & C_{22}^r & \cdots & C_{2n_l}^r \\ \vdots & \vdots & \ddots & \vdots \\ C_{n_l1}^r & C_{n_l2}^r & \cdots & C_{n_ln_l}^r \end{pmatrix} + i \begin{pmatrix} 0 & C_{12}^i & \cdots & C_{1n_l}^i \\ -C_{21}^i & 0 & \cdots & C_{2n_l}^i \\ \vdots & \vdots & \ddots & \vdots \\ -C_{n_l1}^i & -C_{n_l2}^i & \cdots & 0 \end{pmatrix} \\
&= Re(K_W) + iIm(K_W),
\end{aligned}
$$

where $C_m^r(\phi_i, \phi_j) = C_{ij}^r$ and $C_m^i(\phi_i, \phi_j) = C_{ij}^i$.

Now,

$$
K_V = E(\underset{\sim}{V}_m \underset{\sim}{V}_m^*) = E(\underset{\sim}{V}_m \underset{\sim}{V}_m^T)
$$

In order to generate K_V for n_l-tuple case, we reorganize the vector $\underset{\sim}{V}_m$ into the following form.

$$
\begin{aligned}
\underset{\sim}{V}_m &= (W_m^r(\phi_1), \cdots, W_m^r(\phi_{n_l}), W_m^i(\phi_1), \cdots, W_m^r(\phi_{n_l}))^T \\
&= (Re(\underset{\sim}{W}_m), Im(\underset{\sim}{W}_m))^T
\end{aligned}
$$

that is, we grouped all real components and imaginary components together. Hence,

$$
\begin{aligned}
K_V &= E(\underset{\sim}{V}_m \underset{\sim}{V}_m^T) \\
&= \begin{pmatrix} E[Re(\underset{\sim}{W}_m)Re(\underset{\sim}{W}_m)^T] & E[Re(\underset{\sim}{W}_m)Im(\underset{\sim}{W}_m)^T] \\ E[Im(\underset{\sim}{W}_m)Re(\underset{\sim}{W}_m)^T] & E[Im(\underset{\sim}{W}_m)Im(\underset{\sim}{W}_m)^T] \end{pmatrix}_{2n_l \times 2n_l}
\end{aligned}
$$

Since $\underset{\sim}{W}_m$ is circularly-symmetric from 1.15 we can get the following results,

$$
E[Re(\underset{\sim}{W}_m)Re(\underset{\sim}{W}_m)^T] = E[Im(\underset{\sim}{W}_m)Im(\underset{\sim}{W}_m)^T] = \frac{1}{2}(Re(K_W))_{n_l \times n_l}
$$

$$
E[Re(\underset{\sim}{W}_m)Im(\underset{\sim}{W}_m)^T] = -E[Im(\underset{\sim}{W}_m)Re(\underset{\sim}{W}_m)^T] = \frac{1}{2}(Im(K_W))_{n_l \times n_l}
$$

$$K_V = \frac{1}{2}\begin{pmatrix} Re(K_W) & Im(K_W)^T \\ Im(K_W) & Re(K_W) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} Re(K_W) & -Im(K_W) \\ Im(K_W) & Re(K_W) \end{pmatrix}$$

Since K_V is a non-negative definite and matrix, it can be represented as follows,

$$K_V = Q\Lambda Q^T,$$

where Λ is a diagonal matrix with eigen values (real-positive) of K_V and Q are the corresponding orthonormal eigenvectors. We can choose $A = Q\Lambda^{1/2}Q^T$ to obtain,

$$\underset{\sim}{V}_m = A_{2n_l \times 2n_l} Z_{2n_l \times 1},$$

where $Z = \{z_1, z_2, \ldots, z_{n_l}, z_1^*, z_2^*, \ldots, z_{n_l}^*\}$ and each $z_i \sim N(0,1)$ hence we can get $\underset{\sim}{W}_m$. Now for each latitude $\phi_l, l = 1, 2, \cdots, n_l$ and $\lambda_k, k = 1, 2, \cdots, n_L$ $(N = n_L/2)$, we denote the axially symmetric data (real) as $X(\phi_l, \lambda_k)$. These random variates can be obtained from the equation (5.7), let's rewrite the equation as follows,

$$X(\phi_l, \lambda_k) = W_0(\phi_l) + 2\sum_{m=1}^{N}\left[W_m^r(\phi_l)\cos(m\lambda_k) - W_m^i(\phi_l)\sin(m\lambda_k)\right] \quad (5.8)$$

Remark 6. For the above decomposition the eigen values and eigen vectors of K_V is required and the computational cost is $\mathcal{O}((2n_l)^2))$. However according to Gallenger [Gal08] one could use K_W to get the eigen values and eigen vectors where the cost is $\mathcal{O}(n_l^2)$.

Algorithm 5.1 (Pseudo-code)**.**

- *Choose a cross covariance function,* $R(P,Q)$
- *Initialize the parameters* (C_1, C_2, a, u, p) *and choose a resolution* $\phi_1, \ldots, \phi_{n_l}, \lambda_1, \ldots, \lambda_{n_L}$ *(or* $n_l \times n_L$*),*
- *Derive* $C_m(\phi_P, \phi_Q)$ *based on* $R(P,Q)$ *where* $m = 0, 1, \ldots, n_L/2$*,*
  (1) *for each* m *get* $Re(K_W)$ *and* $Im(K_W)$ *hence obtain* K_V
  (2) *use SVD to get* $\underset{\sim}{V}_m$ *(*$n_l -$ *tuples)*
  (3) *get* $\underset{\sim}{W}_m$*'s from* $\underset{\sim}{V}_m$
- *apply the equation* (5.8) *to generate grid data.*

5.2 Simulation setup

In chapter 4 we discussed about the affects of each parameter and for the simulations we fixed the parameters as follows,

	Parameter values
set 1 :	$C_1 = 1, C_2 = 1, a = 1, u = 1,\ p = 0.5$
set 2 :	$C_1 = 1, C_2 = 2, a = 3, u = 1,\ p = 0.6$

Table 3. Parameter values

to generate data on a sphere. The parameter u is the location parameter and $u = 0$ provides a longitudinally reversible process on a sphere. It is somewhat difficult to model the covariance structure in other terms to generate data when it is closer to Earth's pole (the complexity can be observed in MSU and TOMS data). Therefore we generated the data on $[0, 2\pi/3] \times [0, 2\pi]$ (equivalent to $[-\pi/3, \pi/3] \times [-\pi, \pi]$) grid, with a grid resolution of $1^0 \times 2^0$ ($i.e$ $n_l = 120, n_L = 180 \Rightarrow 21600$ spatial points).

The $C_m(\phi_P, \phi_Q)$ function for all three proposed models as follows,

Model	c_m	Parameters
model 1	: $c_m = Cp^m$ and $c_0 = C$	$m = 0, \pm 1, \pm 2, ...$ $p \in (0, 1)$
model 2	: $c_m = \frac{Cp^m}{m}$ and $c_0 = 0$	$m = \pm 1, \pm 2, ...$ $p \in (0, 1)$
model 3	: $c_m = \frac{C}{m^4}$ and $c_0 = 0$	$m = \pm 1, \pm 2, ...$

Table 4. The c_m functions for covariance models used for data generation

Deriving $\mathbf{C_m}$ for model 1

In our approach it is very crucial to derive $C_m(\phi_P, \phi_Q)$ based on a known $R(P, Q)$, the detailed steps to derive $C_m(\phi_P, \phi_Q)$ from model 1 (4.6) is given below,

$$R(P, Q) = R(\phi_P, \phi_Q, \Delta\lambda) = \tilde{C}(\phi_P, \phi_Q) \frac{1 - p^2}{1 - 2p\cos(\Theta) + p^2},$$

where $\Theta = \Delta\lambda + u(\phi_P - \phi_Q)$, with some choice of C_1, C_2, a, u, and p.

Now,

$$
\begin{aligned}
C_m(\phi_P, \phi_Q) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} R(\phi_P, \phi_Q, \Delta\lambda) e^{-im\Delta\lambda} d\Delta\lambda \\
&= \tilde{C}(\phi_P, \phi_Q) \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1-p^2}{1-2p\cos(\Theta + p^2} e^{-im\Delta\lambda} d\Delta\lambda
\end{aligned}
$$

Next we focus on the following integration.

$$
\int_{-\pi}^{\pi} \frac{1-p^2}{1-2p\cos(x+b)+p^2} e^{-imx} dx,
$$

where we set $x = \Delta\lambda$ and $b = u(\phi_P - \phi_Q)$ and we have,

$$
\begin{aligned}
\frac{1-p^2}{1-2p\cos(x+b)+p^2} &= \frac{2-2p\cos(x+b)-(1-2p\cos(x+b)+p^2)}{1-2p\cos(x+b)+p^2} \\
&= 2 \times \frac{1-p\cos(x+b)}{1-2p\cos(x+b)+p^2} - 1 \\
&= 2 \times \sum_{n=0}^{\infty} p^n \cos n(x+b) - 1 \\
&= 1 + 2\sum_{n=1}^{\infty} p^n (\cos nx \cos(nb) - \sin(nx)\sin(nb)).
\end{aligned}
$$

Therefore, for $m \neq 0$,

$$
\begin{aligned}
& \int_{-\pi}^{\pi} \frac{1-p^2}{1-2p\cos(x+b)+p^2} e^{-imx} dx \\
= \quad & \int_{-\pi}^{\pi} \left[1 + 2\sum_{n=1}^{\infty} p^n (\cos nx \cos(nb) - \sin(nx)\sin(nb)) \right] e^{-imx} dx \\
= \quad & \int_{-\pi}^{\pi} e^{-imx} dx + 2\sum_{n=1}^{\infty} p^n \int_{-\pi}^{\pi} \left[\cos nx \cos(nb) - \sin(nx)\sin(nb)\right] e^{-imx} dx \\
= \quad & 2\sum_{n=1}^{\infty} p^n \left[\cos(nb) \int_{-\pi}^{\pi} \cos(nx) e^{-imx} dx - \sin(nb) \int_{-\pi}^{\pi} \sin(nx) e^{-imx} dx \right] \\
= \quad & 2\sum_{n=1}^{\infty} p^n \left[\pi \cos(nb) \delta(n,m) + \pi i \sin(nb) \right] \\
= \quad & 2\pi p^m e^{imb}
\end{aligned}
$$

That is, for $m \neq 0$,

$$
\begin{aligned}
C_m(\phi_P, \phi_Q) &= \tilde{C}(\phi_P, \phi_Q) \frac{1}{2\pi} (2\pi p^m e^{imb}) \\
&= \tilde{C}(\phi_P, \phi_Q) p^m e^{imb}.
\end{aligned}
$$

And for $m = 0$, $C_0(\phi_P, \phi_Q) = \tilde{C}(\phi_P, \phi_Q)$.

In summary,

$$
C_m(\phi_P, \phi_Q) = \begin{cases} \tilde{C}(\phi_P, \phi_Q), & m = 0 \\ \tilde{C}(\phi_P, \phi_Q) p^m e^{imb}, & m \neq 0. \end{cases}
$$

If the process is longitudinally reversible one can set $u = 0$ suppose $\tilde{C}(\phi_P, \phi_Q)$ given by 4.9 then $C_m(\phi_P, \phi_Q)$ for model1,

$$C_m(\phi_P, \phi_Q) = \begin{cases} C_1 \left(C_2 - e^{-a|\phi_P|} - e^{-a|\phi_Q|} + e^{-a|\phi_P - \phi_Q|} \right) & m = 0 \\ C_1 \left(C_2 - e^{-a|\phi_P|} - e^{-a|\phi_Q|} + e^{-a|\phi_P - \phi_Q|} \right) p^m & m \neq 0. \end{cases}$$

The model 1 yields a non zero mean ($c_0 \neq 0$) random process on a sphere, if the process is non-zero mean we showed that (in chapter 4) cross covariance is bias with a bias c_0. In contrast,model 2 and 3 yields a zero mean process on a sphere.

In contrast to the proposed C_m approach one could directly use the covariance matrix ($i.e.R(P,Q)$) to generate data on $\mathbb{S}^2$. The covariance matrix is a real block circuant matix as will take the following form,

$$R(P,Q) = \begin{bmatrix} R_0 & R_1 & R_2 & \cdots & R_{n_L-1} \\ R_{n_L-1} & R_0 & R_1 & \cdots & R_{n_L-2} \\ R_{n_L-2} & R_{n_L-1} & R_0 & \cdots & R_{n_L-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_1 & R_2 & R_3 & \cdots & R_0 \end{bmatrix}_{n_l n_L \times n_l n_L} \tag{5.9}$$

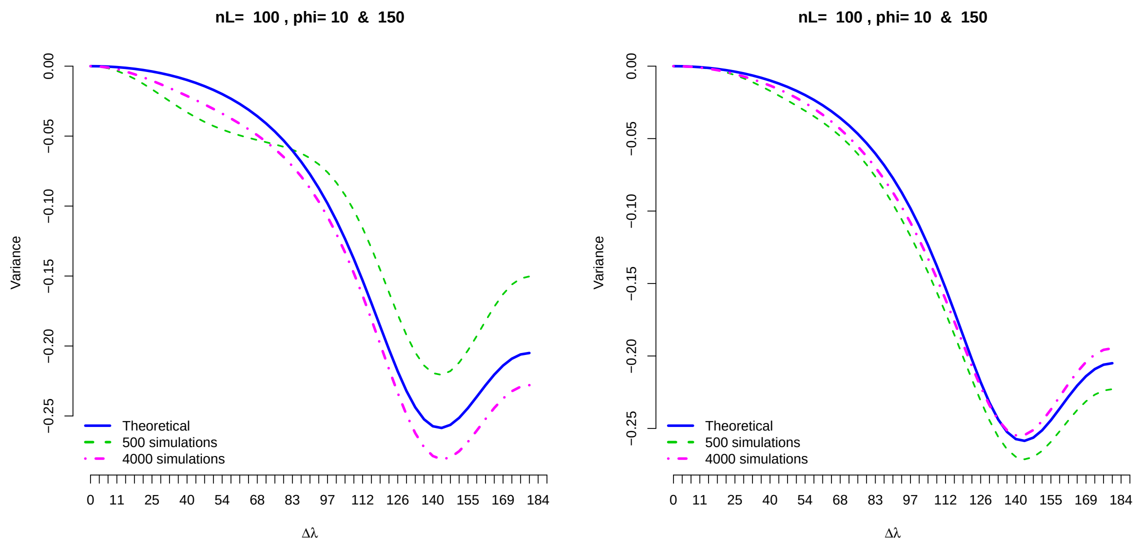
where R_j's are $n_l \times n_l$ sub-matrices of real-valued elements; here R_0 is the covariance matrix between latitudes at logitude 1, R_1 is the covariance matrix between latitudes at logitude 1 and longitude 2, and so on. For the data generation process one would need to compute $R(P,Q)^{1/2}$ in order to use algorithm 5.1. [Li13] pointed out some approaches to find the eigen values using the properties of block circulant matrices *i.e.* all eigen values are related to $R_0, R_1, \ldots, R_{n_L-1}$ sub matrices (each R_j provides n_l eigen values). However, these eigen values to could be real or complex-valued. [Tee05] proposed some methods to find the eigen values when sub matirices are symmetric. It is somewhat unclear to find eigen values and eigen vectors (to obtain $R(P,Q)^{1/2}$)) of a block circulant matrix when the sub matrices are not symmetric.

5.3 Results

To our knowledge, currently there are no methods developed to test axially symmetry in global data, in fact this could be another research opportunity in spatial statistics. Therefore, to validate the compatibility of generated axially symmetric global data we compared the MOM estimate to its theoretical value more preciously the cross variogram estimator. The cross variogram estimator (4.13) is unbiased and in the case of circle we showed that variogram estimator is inconsistent and we expect a similar result in the case of sphere (the proof is left out as future work). We consider multiple to compare the simulated data at multiple pairs of latitudes with fixed number of longitudes ($n_L = 100$). The cross variogram estimator is almost identical to it's theoretical value if pairs (latitudes) were closer. Therefore we will demonstrate a case with larger latitude difference ($\phi_P = 10^0, \phi_Q = 150^0$ equivalent to 70^0S and 60^0N) in which will capture the largest possible errors.

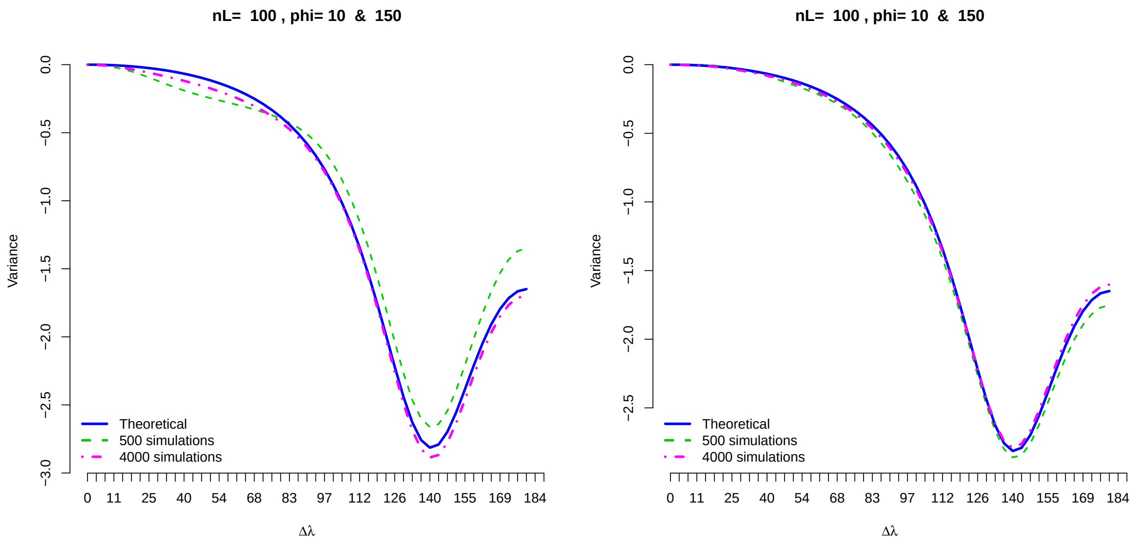
Comparison of MOM estimators

Now we compare the cross variogram estimator given in (4.13), when using $R(P,Q)$ directly compared to proposed C_m approach. Higher dimension is the biggest draw back when using $R(P,Q)$ directly and in some occasions $R(P,Q)$ will provide negative eigen values for example model 3. Hence one cannot perform SVD to generate data. In contrast, it is guaranteed that $C_m(\phi_P, \phi_Q)$ is positive definite and Hermitian. However, it might be a challenge to derive $C_m(\phi_P, \phi_Q)$ from a given $R(P,Q)$ model.



(a) Using $R(P,Q)$ (b) Using C_m

Figure 12. Using parameter set 1 to compare variogram estimator for model 1, solid line (blue) the theoretical values of cross variogram and dashed lines (green, purple) represents the estimates for 500 and 4000 simulations respectively.



(a) Using $R(P, Q)$ (b) Using C_m

Figure 13. Cross variogram comparison using parameter set 2 similar to 12

Results for longitudinally reversible processes

The parameter $u = 0$ yields a longitudinally reversible processes on a sphere (see Figure 11 1(d)) in all proposed models. In all three models the cross variogram estimator converges ($a.s.$) to its theoretical value ($<$ 500 simulations).

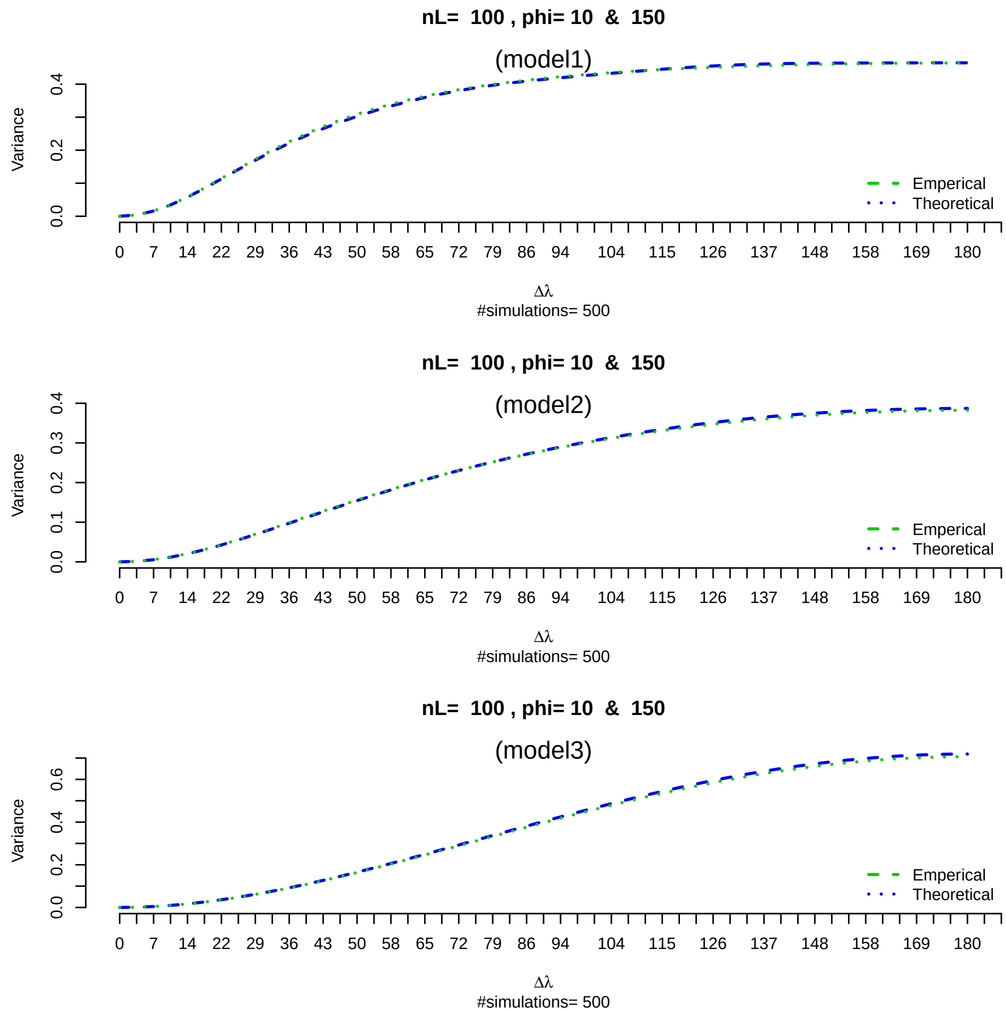


Figure 14. The cross variogram estimator comparison for longitudinally reversible process using model 1, model 2 and model 3 (when $u = 0$).

Comparison of cross covarince

When the random process on a sphere is non zero mean, the cross covariance estimator is biased ($c_0 \neq 0$). Therefore we compared the cross covariance estimator given by 4.11 for zero mean processes (model 2, model 3) on a sphere. In order to compare the cross covariance we used two pairs of latitudes, $\phi = 70, 80$ (20^0S

,10^0S) and $\phi = 60, 120$ (30^0S, 40^0N). The cross covariance estimate match with the theoretical value

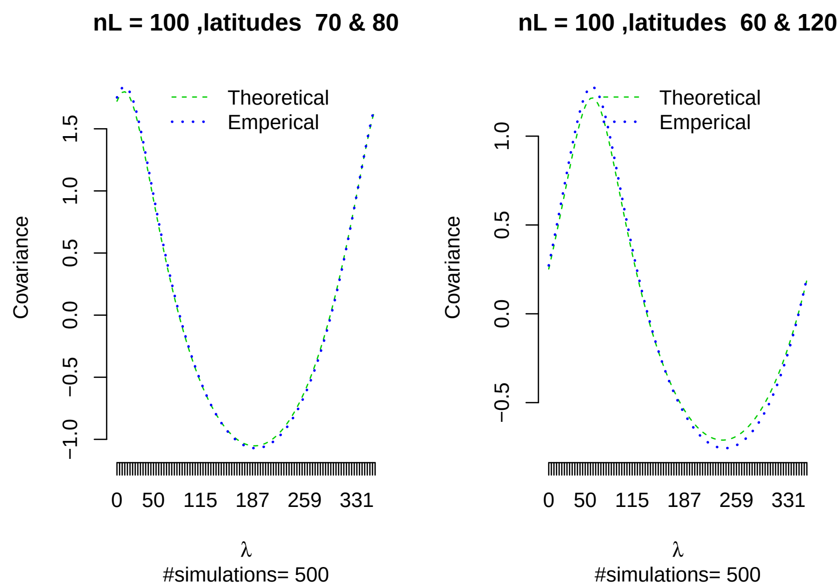


Figure 15. Cross covariance comparison of model 2

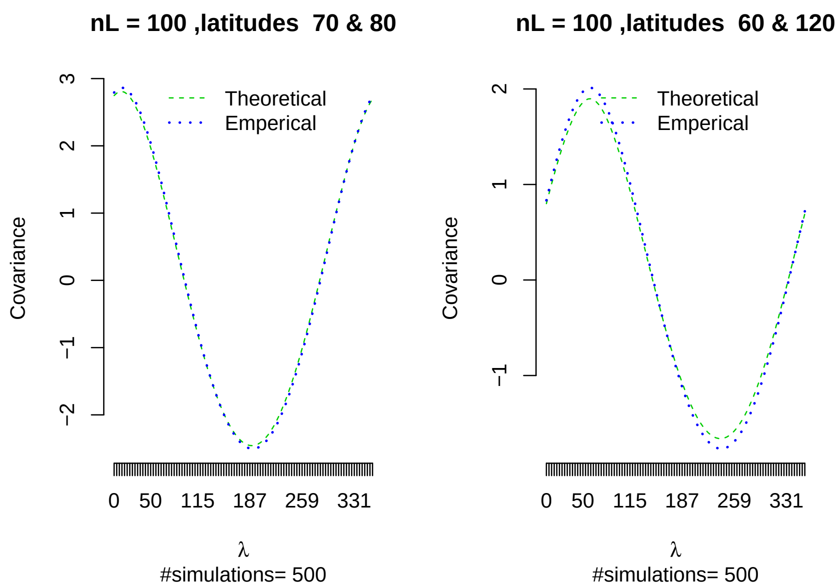


Figure 16. Cross covariance comparison of model 3

The proposed covariance models are functions of latitude difference and the cross covariance estimator converges to its theoretical value with fewer simulations when the difference is small compared to larger latitude differences (the slow convergence at inflection points).

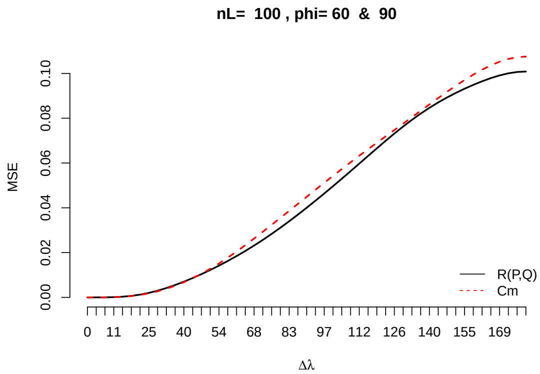
Comparison of MSE

The mean square error (MSE) of cross variogram was computed for both C_m and direct $R(P, Q)$ approaches. In order to compute the MSE multiple pairs of latitudes where considered,

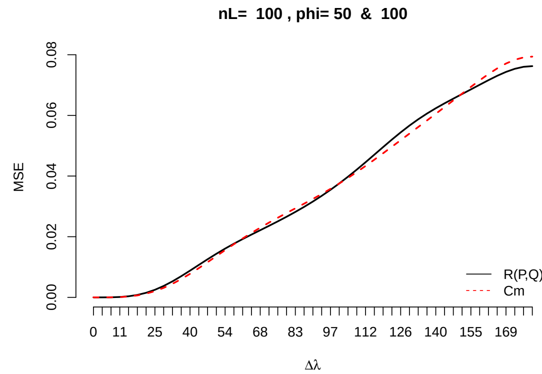
$$
\begin{aligned}
MSE &= \frac{1}{n_L}\sum(var + bias^2) \\
&= \frac{1}{n_L}\sum_{j=1}^{n_L}\left[\frac{1}{nn}\sum_{i=1}^{nn}\left(\hat{\gamma}_i(j\Delta\lambda) - \overline{\hat{\gamma}(j\Delta\lambda)})^2\right) + (\gamma_(j\Delta\lambda) - \overline{\hat{\gamma}(j\Delta\lambda)})^2\right]
\end{aligned}
$$

		Set 1		Set 2	
Model	(ϕ_P, ϕ_Q)	$R(P, Q)$	C_m	$R(P, Q)$	C_m
Model1	(60, 90)	2.300	2.427	–	–
	(50, 100)	1.784	1.782	–	–
	(10, 150)	0.564	0.623	–	–
Model2	(60, 90)	2.000	2.080	12.452	13.021
	(50, 100)	1.437	1.459	9.196	9.262
	(10, 150)	0.457	0.512	6.034	7.266

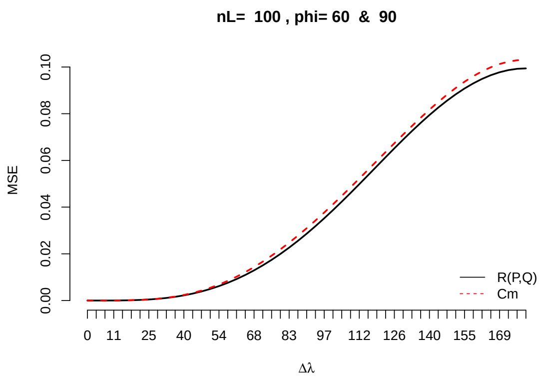
Table 5. MSE comparison for C_m and $R(P, Q)$ approaches, Set 1 and Set 2 are referring to the set of parameters discussed in simulation setup



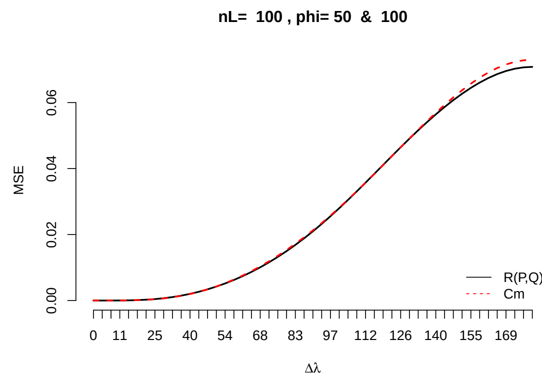
(a) Model 1 (pair 1)



(b) Model 1 (pair 2)



(c) Model 2 (pair 1)



(d) Model 2 (pair 2)

Figure 17. MSE comparison between C_m and $R(P,Q)$ using 500 simulations; pair 1 $(30^0 S, 0^0)$, pair 2 $(40^0 S, 10^0 N)$ figures (a) - (b) is the comparison for model 1 and figure (c)-(d) is the comparison for model 2

Based on 500 simulations the MSE for C_m is closer to its counter part, $R(P,Q)$. However, one should consider the huge reduction in dimension when using the proposed C_m approach. In the data generation process it is not computationaly feasible to perform a SVD with large dimensions. Moreover, obtaining $R(P,Q)^{1/2}$ with the light of block circulant matrix properties seems unclear as we cannot obtain real-valued eigen values for proposed model 3 (4.8). In our approach MSE slightly higher (almost similar) higher, high variance but low bias compared to direct R(P, Q).

Generated data

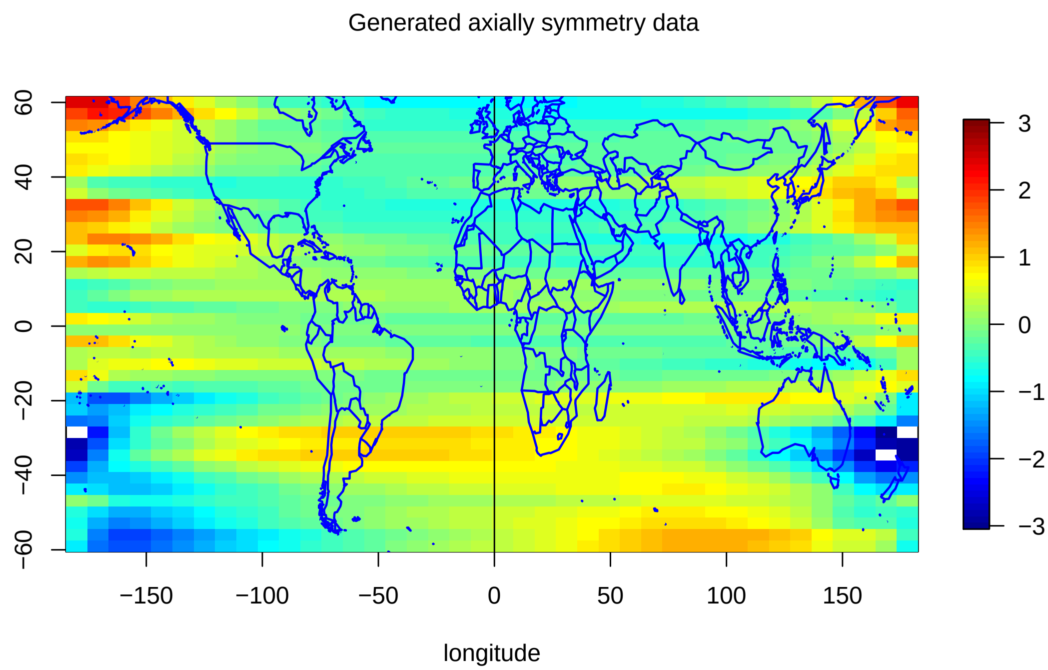


Figure 18. A snapshot of global data generated based on C_m approach using zero mean random process (model 2)

The above figure is a snapshot of the global data generated based on model 2 and are some what compliance with practical geo-spatial data (MSU and TOMS), clearly there are spatial trends within the latitudes but not within longitudes. However we observed some inconsistencies (strong spots) closer to the boundary points of longitudes ($\lambda \to \pm\pi$).

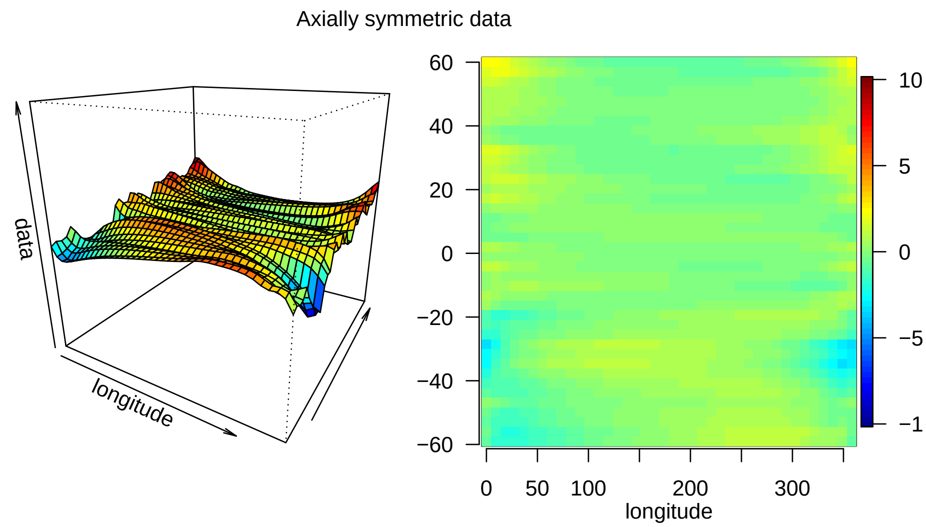


Figure 19. One snapshot of the axially symmetric data generated based on model 2, grid resolution $2^0 \times 1^0$ (data scale -10 and 10).

The above figure refers to the data snapshot given by figure (18) and it is clearly evident that trends are within latitudes not within longitudes. Four snapshots of the gridded data generated based on all models are given in appendix A.

CHAPTER VI

FUTURE RESEARCH

In Chapter 3 we proved in general the MOM variogram estimator is inconsistent in the case of circle, but we continue to investigate further on this result and explore if there are any special ergodic condtions, parallel to the ones in $\mathbb{R}^d$, that MOM variogram estimator might be consistent. Further we suspect that MOM cross variogram estimator is also inconsistent in the case of $\mathbb{S}^2$ and we wish to explore more details about the estimator and evaluate if it is consistent under any restircted conditions. Then we wish to propose more robust variogram estimators, similar to [**?**], that are useful in the case of $\mathbb{S}^2$.

In the global data generation process we fixed five paramters based on convenince (under the given bounds) but we wish to optimize these paramters based on real data or fix these paramters which are more practially useful for geo-scientists. In order to use the proposed algorithm 5.1 for other covariance models (for example mat$\acute{e}$rn class) that are valid on $\mathbb{S}^2$, first one would need to obtain C_m form a given $R(P, Q)$ but this might be challenging. We wish to explore alternative techniques such as Discrete Fourier Transformations (DFT) to obtain C_m and compare the robustness of other covariance models.

In spatial statistics, when analyzing or modelling spatial data it is very important to make predictions at an unobserved location(s). Kriging is a widely used technique to make predictions in geo-statics but the issue is such it could only make local (Euclidean space) prediction. So far there are no reliable methods developed to make global predictions, [Ste99] proposed a parametric model with 170 parameters to estimate, [CJ08] proposed a model with 396 parameters to estimate and [Li13] proposed some ideas about global predictions. We wish to enhance the kriging techniques and make use of proposed global data generation methods to make global predictions with less parameters to estimate.

REFERENCES

[BL11] David Bolin and Finn Lindgren, *Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping*, The Annals of Applied Statistics (2011), 523–550.

[CJ08] Noel Cressie and Gardar Johannesson, *Fixed rank kriging for very large spatial data sets*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **70** (2008), no. 1, 209–226.

[Cre93] Noel Cressie, *Statistics for spatial data*, John Wiley & Sons, 1993.

[CSB00] John R Christy, Roy W Spencer, and William D Braswell, *Msu tropospheric temperatures: Dataset construction and radiosonde comparisons*, Journal of Atmospheric and Oceanic Technology **17** (2000), no. 9, 1153–1170.

[DMG83] T De Mazancourt and D Gerlic, *The inverse of a block-circulant matrix*, IEEE transactions on antennas and propagation **31** (1983), no. 5, 808–810.

[DR76] Jean-Marie Dufour and Roch Roy, *On spectral estimation for a homogeneous random process on the circle*, Stochastic Processes and their Applications **4** (1976), no. 2, 107 – 120.

[Gal08] Robert G Gallager, *Circularly-symmetric gaussian random vectors*, preprint (2008), 1–9.

[Gne13] Tilmann Gneiting, *Strictly and non-strictly positive definite functions on spheres*, Bernoulli **19** (2013), no. 4, 1327–1349.

[HS12] Marcin Hitczenko and Michael L Stein, *Some theory for anisotropic processes on the sphere*, Statistical Methodology **9** (2012), no. 1, 211–227.

[HZR11] Chunfeng Huang, Haimeng Zhang, and Scott M Robeson, *On the validity of commonly used covariance and variogram functions on the sphere*, Mathematical Geosciences **43** (2011), no. 6, 721–733.

[HZR12] ______, *A simplified representation of the covariance structure of axially symmetric processes on the sphere*, Statistics & Probability Letters **82** (2012), no. 7, 1346–1351.

[Jon63] Richard H Jones, *Stochastic processes on a sphere*, The Annals of mathematical statistics **34** (1963), no. 1, 213–218.

[JS07] Mikyoung Jun and Michael L Stein, *An approach to producing space: Time covariance functions on spheres*, Technometrics **49** (2007), no. 4, 468–479.

[JS08] ______, *Nonstationary covariance models for global data*, The Annals of Applied Statistics (2008), 1271–1289.

[Lea67] H Cramer-MR Leadbetter, *Stationary and related stochastic processes*, New York (1967).

[Li13] Yang Li, *Non-parametric and semi-parametric estimation of spatial covariance function.*

[LN97] Ta-Hsin Li and Gerald R North, *Aliasing effects and sampling theorems of spherical random fields when sampled on a finite grid*, Annals of the Institute of Statistical Mathematics **49** (1997), no. 2, 341–354.

[LRL11] Finn Lindgren, Håvard Rue, and Johan Lindström, *An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **73** (2011), no. 4, 423–498.

[Mui20] Sir Thomas Muir, *The theory of determinants*, 1920.

[Roy72] Roch Roy, *Spectral analysis for a random process on the circle*, Journal of Applied Probability (1972), 745–757.

[Sch42] I. J. Schoenberg, *Positive definite functions on spheres*, Duke Math. J. **9** (1942), no. 1, 96–108.

[Ste99] Michael L Stein, *Interpolation of spatial data: some theory for kriging*, Springer Science & Business Media, 1999.

[Ste07] ______, *Spatial variation of total column ozone on a global scale*, The Annals of Applied Statistics (2007), 191–210.

[Tee05] Garry J Tee, *Eigenvectors of block circulant and alternating circulant matrices.*

[Tra73] George E Trapp, *Inverses of circulant matrices and block circulant matrices*, Kyungpook Mathematical Journal **13** (1973), no. 1, 11–20.

[Wac13] Hans Wackernagel, *Multivariate geostatistics: an introduction with applications*, Springer Science & Business Media, 2013.

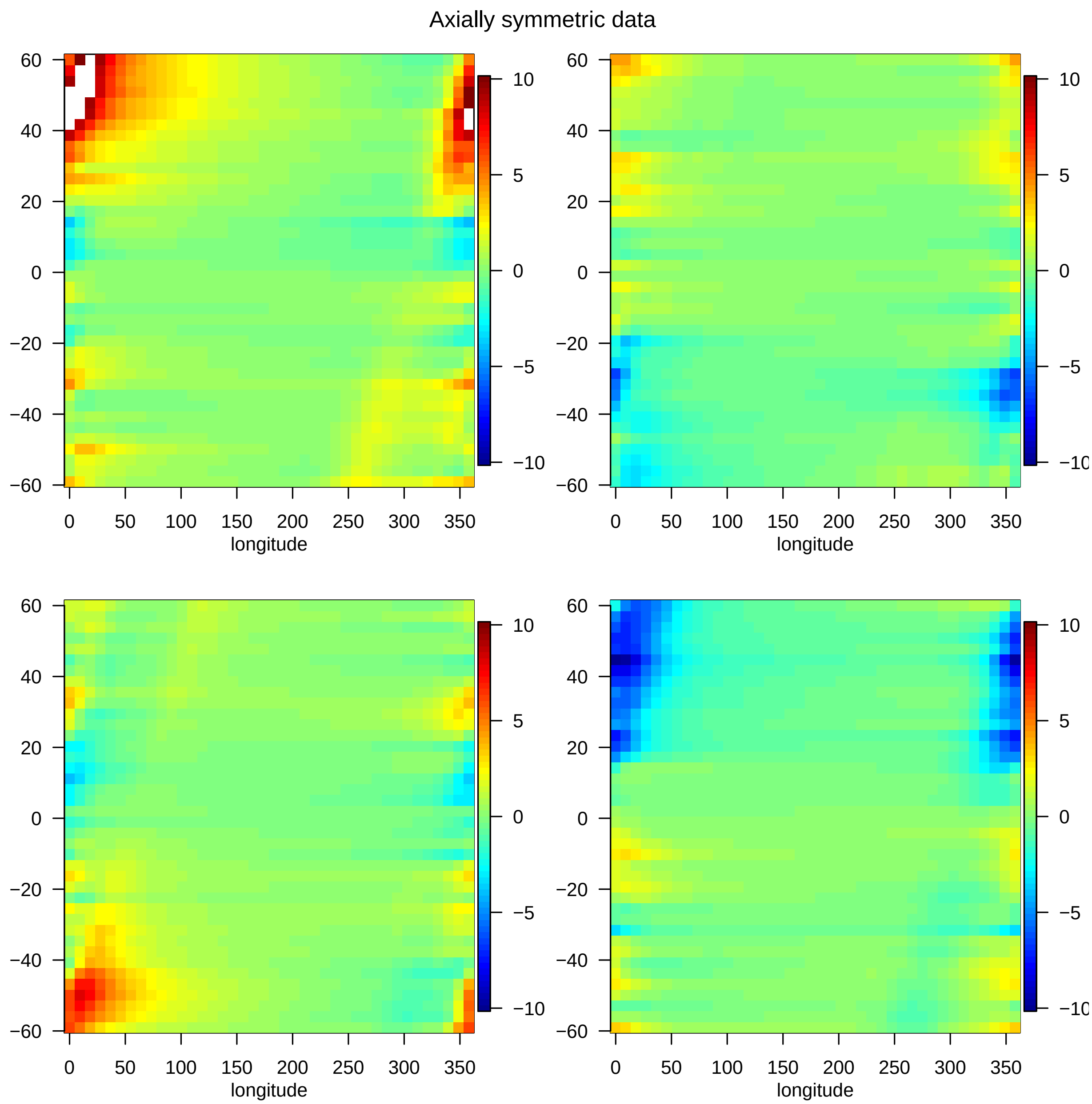
[Yad83] Mikhail Iosifovich Yadrenko, *Spectral theory of random fields*, Optimization Software, 1983.

[Yag61] Akiva M Yaglom, *Second-order homogeneous random fields*, Proc. 4th Berkeley Sympos. Math. Statist. and Prob, vol. 2, 1961, pp. 593–622.

[Yag87] AM Yaglom, *Correlation theory of stationary and related random functions: Vol.: 1: Basic results*, Springer-Verlag, 1987.

APPENDIX A

DATA SNAPSHOTS FOR ALL MODELS



Four consecutive axially symmetric data snapshots based on model 1, grid resolution $2^0 \times 1^0$ (data scale -10 and 10).

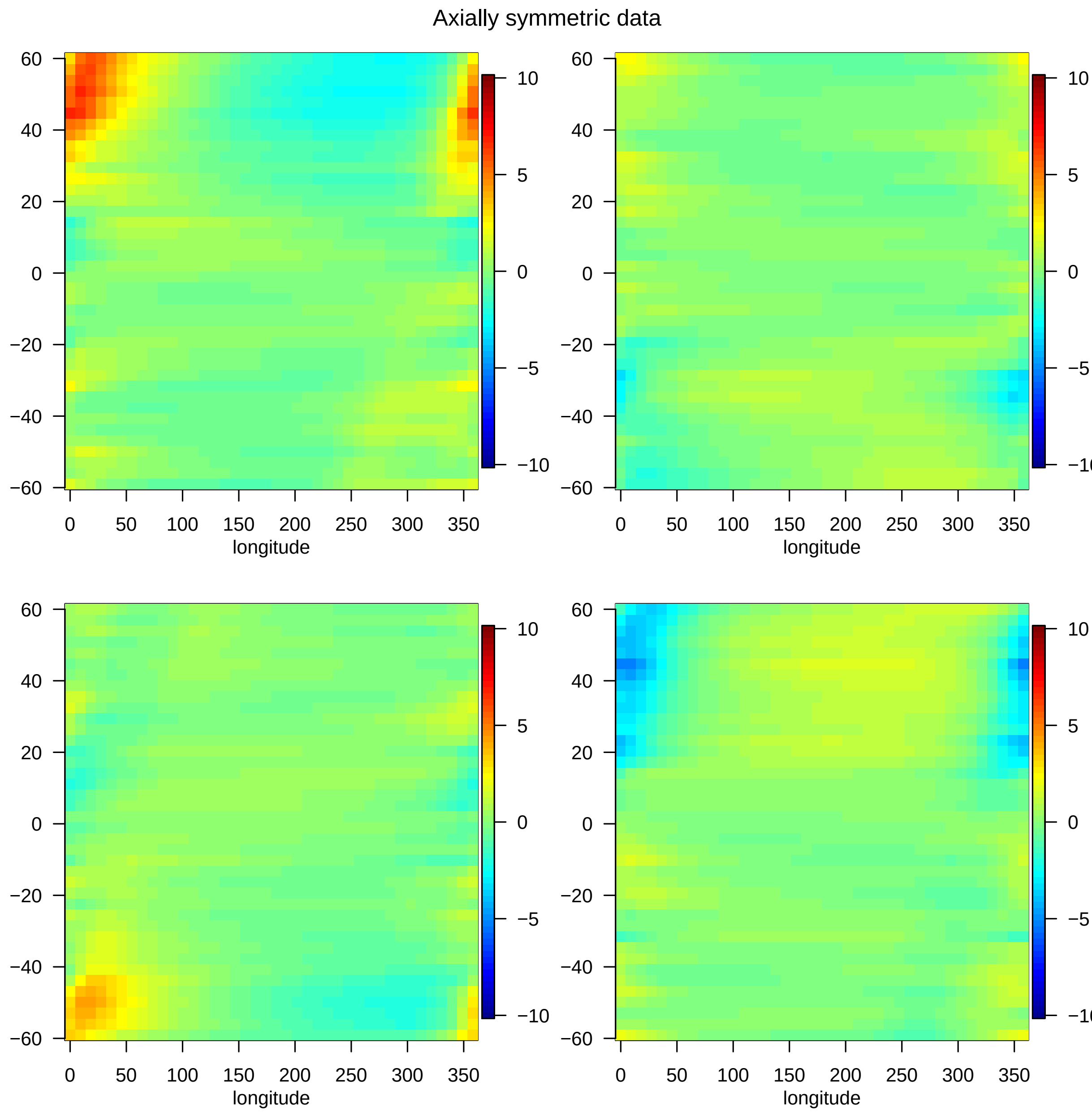
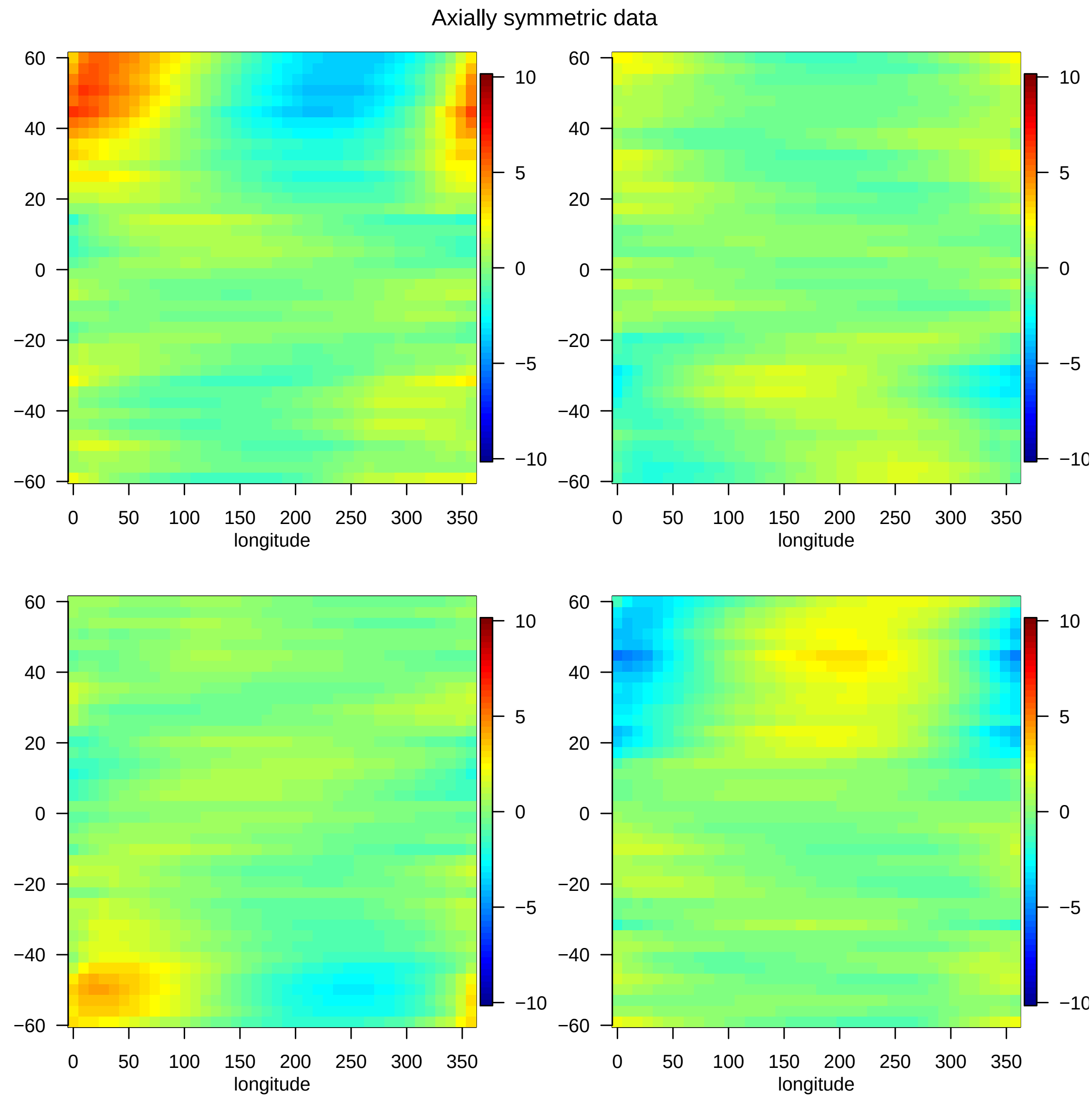


Figure 20. Four consecutive axially symmetric data snapshots based on model 2, grid resolution $2^0 \times 1^0$ (data scale -10 and 10).



Four consecutive axially symmetric data snapshots based on model 3, grid resolution $2^0 \times 1^0$ (data scale -10 and 10).