VANLANGENBERG, CHRISTOPHER D., Ph.D. Data Generation and Esimation for Axially Symmertic Processes on the Sphere . (2016)
Directed by Dr. Haimeng Zhang. 98 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 father, 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

First, I would like to express my special appreciation and gratitude to my advisor Dr. Haimeng Zhang for his motivation and willingness to share the knowledge he aquired over the years. His door was always opened to me for any questions; without his support and guidance this dissertation would not have been possible. Dr. Zhang's advise was priceless, it helped me not only to grow as a young researcher but also as a teacher, a professional and a leader.

I would like to thank my dissertation committee members, Dr. Sat Gupta, Dr. Scott Richter, Dr. Xiaoli Gao, and Dr. Shan Suthaharan, for been in my committee, sharing their ideas, suggestions and comments to improve this dissertation. Especially, I would like to thank Mathematics and Statistics Department for providing me a Teaching Assitantship to support my studies. Further, I would like to thank Dr. Scott Richter for engaging me in his consulting projects and for providing a Research Assiatantship during summer through Statistical Consulting Center. I wish to thank all faculty, staff and my friends at Mathematics and Statistics Department for all their support and fun time at UNCG.

A special thanks to my mother, sister, relatives and friends for their endless support. My father is not with me today to share my results but I would like to dedicate this dissertation to him for his inspiration and encouragement. Finally, I would like to express appreciation to my beloved wife Lakma Fernando who supported me every possible way in this journey and my two kids Ethmi and Kethmin who keep me busy and for bringing happiness and joy to our family.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

In this chapter we will give a brief introduction to some of the basic concepts in spatial statistics, Specifically, we will discuss stationarity and intrinsic stationarity, covariance and variogram functions and their properties, mean square continuity, 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 $\{Z(x) : x \in \Omega\}$, 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.

- $x \in \Omega = N$, the set of all integers: $Z(x)$ is a time series.
- $x \in \Omega = \mathbb{R}^1$: $Z(x)$ is a random process, commonly referred as a stochastic process.
- $x \in \Omega = \mathbb{R}^d$: $Z(x)$ is a random field or a spatial process if $d > 1$.
- $x \in \Omega = \mathbb{S}^2$: $Z(x)$ is a random process on the sphere.
- $x \in \Omega = \mathbb{R}^d \times \mathbb{R}$: $Z(x)$ 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 $\mathbb{R}^d$, 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 $\underset{\sim}{Z} = (Z(x_1), \ldots, Z(x_n))^T$ 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 $h_1, \ldots, h_n \in \mathbb{R}$.

1.1.1 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,

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

The assumption of strict stationarity is normally too strong 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 stationary) if it has a constant mean and finite second moment as well as a (auto-)covariance

function $C(h)$ that depends only on the spatial distance between the 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, 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) $C(h) = a_1 C_1(h) + a_2 C_2(h)$, $\forall a_1, a_2 \geq 0$;

(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 $x_1, x_2, \ldots, x_N \in D \subseteq \mathbb{R}^d$.

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 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 isotropic if the joint distributions are invariant under all rigid motions. *i.e.,* for any orthogonal $d \times d$ matrix H and any $x \in \mathbb{R}^d$

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

Isotropy assumes that it is not required to distinguish one direction from another for the correlation of a random field $Z(x)$. When describing the correlation between random fields at two locations, the correlation of $Z(x)$ at any two locations is the same as long as these two points are within the same distance.

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

$$2\gamma(h) = Var(Z(x + h) - Z(x)). \quad (1.6)$$

Here $\gamma(h)$ is called the semivariogram. If the variogram function solely depends on the distance between 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 the 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, consider the linear semivariogram function given below,

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

where $a, \sigma \in \mathbb{R}$. Since $\gamma(h) \to \infty$ as $h \to \infty$, the process with the above semivariogram is not stationary and $C(h)$ does not exist.

Parallel to 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 $x_1, x_2, \ldots, x_N \in D \subseteq \mathbb{R}^d$.

1.1.2 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 that $\{X_n\}$ 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.

1.1.3 Spectral Representation of a Random Field

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 $U(x)$ and $V(x)$ are stationary so is $Z(x)$. 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 $x_1, x_2, \ldots, x_n$,

$$\sum_{i,j=1}^{n} c_i \bar{c_j} C(x_i - x_j) \geq 0. \tag{1.8}$$

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 $E(Z_i \bar{Z}_j) = 0$ ($\bar{Z}$ represents the conjugate of the complex number Z), $i \neq j$ and $E|Z_i|^2 = f_i$. Then the random sum

$$Z(x) = \sum_{k=1}^{n} Z_k e^{i\omega_k^T x}, \quad x \in \mathbb{R}^d \tag{1.9}$$

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

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 called the spectral distribution. Here is a more general result from Bochner (Stein, 1999).

Theorem 1.1 (Bochner's Theorem).
A complex valued covariance function $C(\cdot)$ on $\mathbb{R}^d$ is 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 an f exists) we can use the inversion formula to obtain f:

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

1.1.4 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) The Gaussian covariance function is 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) The Matérn class has a more practical use and it is more frequently used in spatial statistics. The spectral density is 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 is 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 ν is, the smoother the Y_ν. Further, Y_ν will be m times mean square differentiable if and only if $\nu > m$. When ν is of the form of $m + 1/2$ with m being a non negative integer, the spectral density is rational and the covariance function is of the form $e^{-\alpha|h|}$. polynomial($|h|$). For example, when $\nu = \frac{1}{2}$, $C(h)$ corresponds to exponential model and when $\nu = \frac{3}{2}$, it is the transformation of the 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

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 proba-

bility distribution for all real ϕ. Since $E[e^{i\phi}Z] = e^{i\phi}E[Z]$, any circularly-symmetric complex random variable must have $E[Z] = 0$.

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 complex random vector of dimension $2n$. Then its covariance matrix $K_{\underset{\sim}{Z}}$ and the pseudo-covariance matrix $M_{\underset{\sim}{Z}}$ are defined as follows:

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

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

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 both Z^{Re} and Z^{Im} are real and jointly Gaussian. Now 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 where each follows $N(0, 1/2)$, so $\underset{\sim}{Z}$ is complex Gaussian.

The covariance matrix defined by (1.13) is given by

$$M_{\underset{\sim}{Z}} = E\begin{bmatrix} Z_1^2 & Z_1Z_1^* \\ Z_1Z_1^* & Z_1^2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

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

$$K_{\underset{\sim}{Z}} = E\begin{bmatrix} Z_1 Z_1^* & Z_1^2 \\ Z_1^2 & Z_1 Z_1^* \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

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.*, $M_{\underset{\sim}{Z}} \equiv K_{\underset{\sim}{Z}}$.

The covariance matrix of the real $\underset{\sim}{Z} = (\underset{\sim}{Z}^{Re}, \underset{\sim}{Z}^{Im})^T$ random vector with dimension of $2n$, 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 follows:

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

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

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

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

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.

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

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 a 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 $\lambda_j = \lambda_{n-j}$ (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 Tee (2005) Hermitian matrices has a full set of orthogonal eigenvectors with corresponding real eigenvalues.

1.3.1 Block Circulant Matrices

The idea of a block circulant matrix was first proposed by Muir (1920). 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. De Mazancourt and Gerlic (1983) 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 an $np \times np$ identity matrix, the eigenvalues of M given by ρ_l, and the eigenmatrix of M can be given by $Q_{np\times np} = \{\underset{\sim}{\psi_0}, \underset{\sim}{\psi_1}, \ldots, \underset{\sim}{\psi_{n-1}}\}$. It can be shown (Trapp, 1973) that $Q^{-1} = Q^*/n$ where Q^* is the conjugate transpose of Q. Thus 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 given by the discrete Fourier transform of the blocks A_j,

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

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 dependent on the dimension of B and the eigenvalues of B (ρ_l's). In other words, B is not dependent 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 it mmediately 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 $Q^{-1} = Q^T$.

CHAPTER II

LITERATURE REVIEW

2.1 Spatial Data

What is meant 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 in the form of maps in 1686 but spatial modeling did not start until 1907 (Cressie, 1993). There are many questions that interest geoscientists and engineers about spatial data. Many questions naturally arise such as how to model a spatial process and then use these models to make predictions at unobserved locations. There are many challenges when modeling spatial data, however. Every point (location observed) is a random variable and only one observation/measurement is available. In addition the number of unknowns to estimate is quite large compared to the available data, which creates a high-dimensional data problem. For example, if data are observed at 10 locations, in order to estimate the variance-covariance matrix to characterize the spatial dependency for future predictions, there will be 55 unknown entities in the variance-covariance matrix to be estimated. We will discuss some basic properties of geospatial data by exploring some popular data sets from the literature.

Since 1978 Microwave Sounding Units (MSU) have measured 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 processed can be found in Christy et al. (2000). Satellites do not measure temperature directly but measure radiances in various wavelength bands and then mathematically invert to obtain the actual temperature.

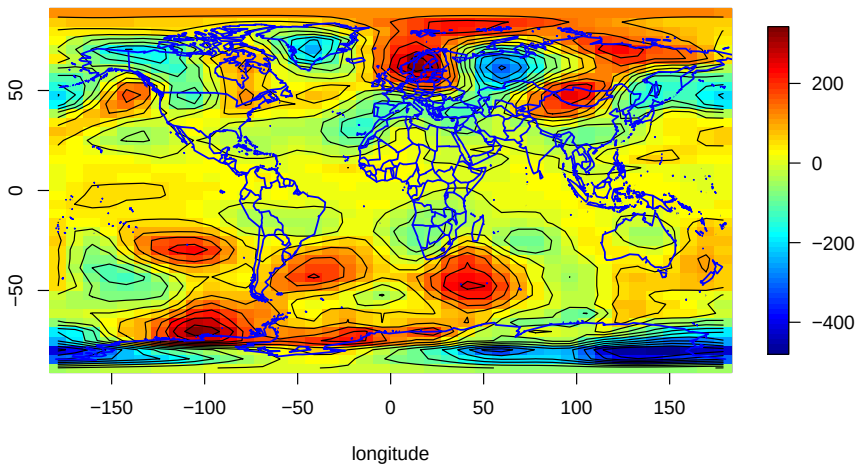
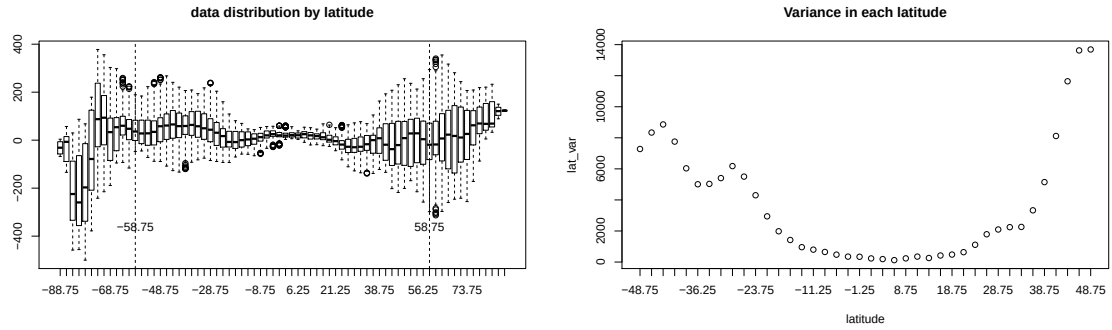


Figure 1. MSU data observed in August 2002, grid resolution 2.5°latitude × 2.5°longitude with 10368 observations.

[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 popular global data set discussed in literature which has more than 20000 spatial points or gridded points (Cressie and Johannesson, 2008; Jun and Stein, 2008; Stein, 2007). There were some missing values in this data set. Stein (2007) 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 axial symmetry (discussed in Chapter 4) of the global data.

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

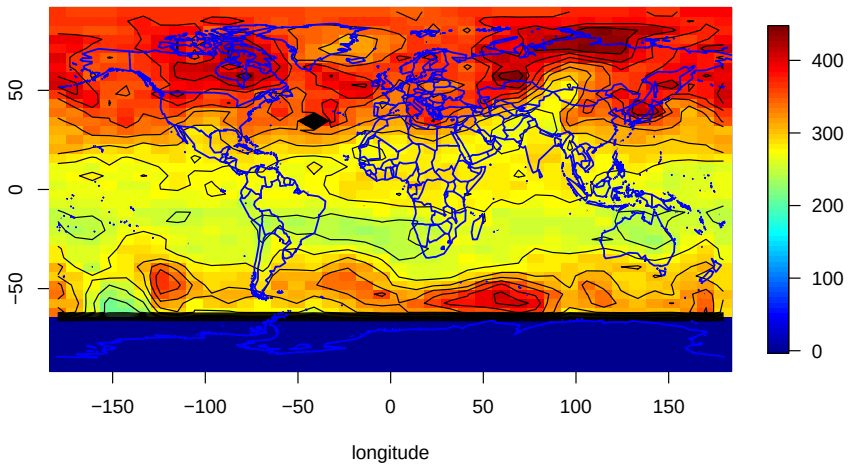
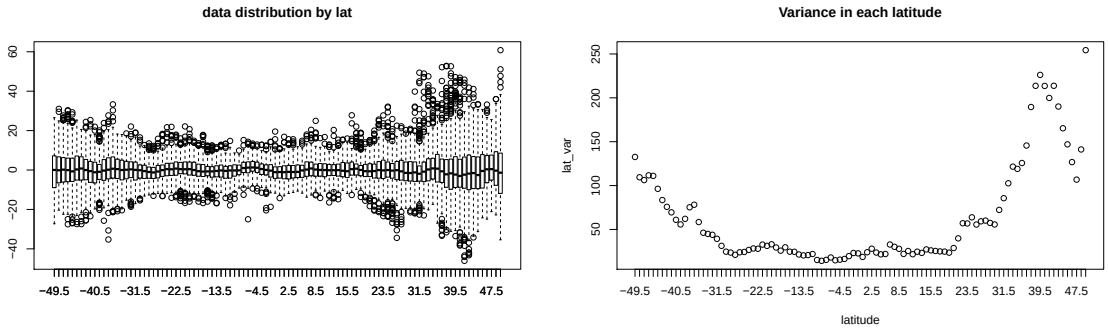


Figure 3. TOMS data: resolution 1°latitude $\times 1.25°$ longitude in May, 1-6 1990. The instrument used back-scattered sunlight, therefore measurements were not available south of $73°S$ during this week.



(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.

2.2 Research Studies in Spatial Data

There has been extensive statistical research on methodologies and techniques developed under Euclidean space $\mathbb{R}^d$. Approaches that are valid in $\mathbb{R}^d$ 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. Table 1 lists some commonly used covariance models that are valid on $\mathbb{R}^3$.

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

Family	$C(h)$	Parameters
Matérn	$\frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}(\frac{h}{\phi})^{\nu}Y_{\nu}(\frac{h}{\phi})$	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

However, this can have unforeseen impacts, such as making use of models that are valid in $\mathbb{R}^d$ but in fact might not be valid under spherical coordinate systems. Huang et al. (2011) have investigated some of commonly used covariance models that are valid in $\mathbb{R}^d$, and they pointed out that many are actually invalid on the sphere. In particular, Gneiting (2013) proved that commonly used Matérn 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 global data. Note that the assumption of homogeneity on the sphere requires that the mean of the random process is constant and that 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, Stein (2007) argued that Total Ozone Mapping Spectrometer (TOMS) data varies strongly with latitudes and thus homogeneous models are not suitable. 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. Monte Carlo Markov Chain (MCMC) modelling is another approach to model non-stationary covariance models on a sphere. Lindgren et al. (2011) 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 Bolin and Lindgren (2011) 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 claimed that the method is more efficient than standard MCMC procedures.

The analysis and modeling of axially symmetric data on the sphere has received increasing attention in literature in recent years. It was first introduced by Jones

(1963), 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 and moisture often exist homogeneity on longitudes rather than latitudes. Stein (2007) used spherical harmonics to model (TOMS) data that exhibit an axial symmetry. Jun and Stein (2008) applied first-order differential operators to an isotropic process to draw conclusions about the local properties of axially symmetric spatio-temporal processes. Hitczenko and Stein (2012) investigated the properties and theory of different forms of axially symmetric processes on the sphere. Li (2013) used convolution methods with Matérn-type kernel functions to capture the non-stationarity of random fields on a sphere. Huang et al. (2012) developed a new and simplified representation for a valid axially symmetric process and also explored the construction of parametric models for axially symmetric processes.

The computational cost for modeling and analyzing axially symmetric data is very expensive. As we will see in Chapter 4, the covariance function for axially symmetric processes requires triple summations, which one is to estimate $O(n^3)$ parameters. Although the covariance structure given by Huang et al. (2012) might potentially reduce the number of parameters to be estimated in the order of $O(n^2)$, the large data sets from global sensors and satellites often add much more computational cost. Stein (2007) used 170 parameters from an axially symmetric covariance structure to model TOMS data but was still not able to capture the global dependency. Cressie and Johannesson (2008) used more than 396 parameters when they modeled global

data. Hence, 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 *–citation–*. 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 that in literature there is a lack of simulations that generate global data that follow the axially symmetric covariance structure. On the other hand, 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 $\mathbb{R}^d$. 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 $\mathbb{R}^d$ 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 for the stationary process on the circle based on Method of Moments (MOM). In contrast to the results given in time series and the Euclidean space, the MOM covariance estimator is biased and the true covariance function might not be identifiable based on the MOM estimator. On the other hand, the MOM variogram estimator is unbiased, but it is inconsistent under the assumption of Gaussianity. 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. Our main focus in this chapter is the axially symmetric process and its covariance function representation through the discrete Fourier transform. The parametric models for characterizing such processes will also be discussed. In particular, we extend the models given in Huang et al. (2012) 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 implement an algorithm to generate axially symmetric data based on the given covariance structure. We validate our generated data via simulations. 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 functions have been commonly used in the 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 $\mathbb{R}^1$ 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 certain ergodic conditions, for example, the covariance function converges to zero as the lag $h \to \infty$. Thus,

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

implying the consistency of $\bar{X}$ when estimating μ (for example Cressie, 1993).

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 $O(1/n)$, the bias for the variogram MOM estimator is smaller. Furthermore, the asymptotic joint Gaussianity of the estimators $\{\hat{C}(h)\}$ and $2\hat{\gamma}(h)$ has also been established under the same ergodic conditions (*i.e.*, conditions that ensure the dependence in the process dies off sufficiently quickly as the 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 Cressie, 1985; Fuller, 2009) as 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 $\mathbb{R}^d$ (Cressie, 1993).

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 $\mathbb{R}^d$ belong to the first one. There have been extensive discussions of the infill asymptotics in the 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, discussion about the asymptotics on the circle and sphere is lacking. In particular, with the increasing interest in the study of global data on the sphere, it is necessary that the infill asymptotics of existing estimators 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 follows. 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 with non-estimable bias, while the variogram MOM estimator is unbiased but inconsistent. Our results are supplemented via simulations.

3.1.1 Random Process on a Circle

Let $X(t)$ be the random process on the unit circle $\mathbb{S}^1$. 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 (Dufour and Roy, 1976),

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

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 $s, t \in \mathbb{S}^1$. 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

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

with $a_n \geq 0$, and $\delta(n, m) = 1$ if $n = m$, and 0 otherwise. Further 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, and the variogram function depends only on the angular distance θ, given by

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

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

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

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

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

3.1.2 Mean and Covariance Estimation on the Circle

We now consider estimation of 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 $\underset{\sim}{X} = (X(t_1), X(t_2), \cdots, X(t_n))^T$ as the observed random vector. When the underlying process $X(t)$ is stationary on the unit circle, the variance-covariance matrix of $\underset{\sim}{X}$ given by

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

where Σ is a symmetric circulant matrix with elements

$$C(0), C(\delta), C(2\delta), \cdots, C((N-1)\delta), C(\pi), C((N-1)\delta), \cdots, C(\delta), \text{ where } \delta = 2\pi/n.$$

Therefore the sample mean (denoting $\underset{\sim}{1}_n = (1, 1, \ldots, 1)^T_{n\times 1}$),

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

is an unbiased estimator of μ with variance given by

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

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 $\mathbb{R}^d$.

Remark 2. Recall that, if $X(t)$ is a 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 $a_0 = 0$ so that $X(t)$ 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 to estimate 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 an unbiased and consistent estimator of a_0.

3.2 Data Generation on a Circle

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

$$C(\theta) = C_1 e^{-a|\theta|} \quad a > 0, C_1 > 0, \theta \in [0, \pi], \tag{3.1}$$

$$C(\theta) = c_0 - (|\theta|/a)^{\alpha} \quad a > 0, \alpha \in (0, 2]. \tag{3.2}$$

where $\theta \in [0, \pi]$ and $c_0 > 0$ satisfies $c_0 \geq \int_0^{\pi}(\theta/a)^{\alpha}\sin\theta d\theta$.

Recall that $\underset{\sim}{X} = (X(t_1), X(t_2), \ldots, X(t_n))^T$ is the observed gridded data on the circle. Its variance-covariance matrix Σ is symmetric and circulant. Hence it can be decomposed as

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

where $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ and $Q = \{\psi_1, \psi_2, \cdots, \psi_n\}$ with $\lambda_i, i = 1, 2 \ldots, n$, are eigenvalues and $\psi_1, \psi_2, \cdots, \psi_n$ are eigenvectors of the circulant matrix, respectively.

(See Section 1.3). Therefore, let $\underset{\sim}{Z}$ be *i.i.d.* standard normal random variates. Then

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

is the observed data vector that follows the given covariance function.

3.3 Covariance MOM Estimator

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

$$\hat{C}(\Delta\lambda) = \frac{1}{n}\sum_{i=1}^{n}(X(t_i + \Delta\lambda) - \bar{X})(X(t_i) - \bar{X}), \tag{3.3}$$

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

For the simulation, we set $C_1 = a = 1$ and choose $\alpha = 0.5$, $c_0 \geq \int_0^{\pi}(\theta)^{0.5}\sin(\theta)d\theta$. From the Fresnel integral, it can be shown that $c_0 \geq 2.4353$. Now we compare the covariance estimator (empirical) to its theoretical covariance given by (3.1) and (3.2), respectively. We computed the MOM estimator $\hat{C}(\theta)$ with $n = 48$ gridded observations on the circle with 500 repetitions.

Remark 4. One can easily notice a shift between the theoretical and empirical values appearing on both graphs (Figure 5). The shift can be shown to be approximately equal to a_0, where $a_0 = \frac{1}{\pi}\int_0^{\pi} C(\theta)d\theta$. For both covariance functions considered, we can obtain

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

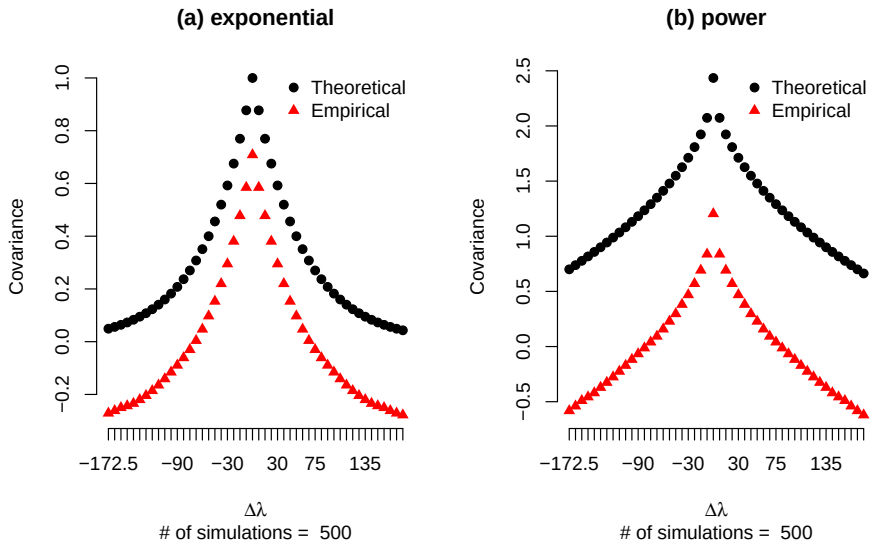


Figure 5. Comparison between theoretical and empirical covariances without any adjustments on a circle.

Next we consider the covariance function $D(\theta)$, after subtracting a_0 from $C(\theta)$. Since the bias seems to be equal to a_0

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

From Remark 2, $D(\theta)$ is now the covariance function of a stationary process on the circle. Where the constant term of its spectral representation equals to zero. The simulation setup is the same as above except that $C(\theta)$ is replaced with $D(\theta)$. The results show that the empirical and theoretical values match very well (Figure 6).

Remark 5. The covariance estimator is biased and the bias is close to a_0, which is non-estimable. However if *i.i.d.* copies of random data on the same circle are available then one can estimate a_0 *i.e.*, $\hat{a} = var(\bar{X})$. The new covariance estimator could then

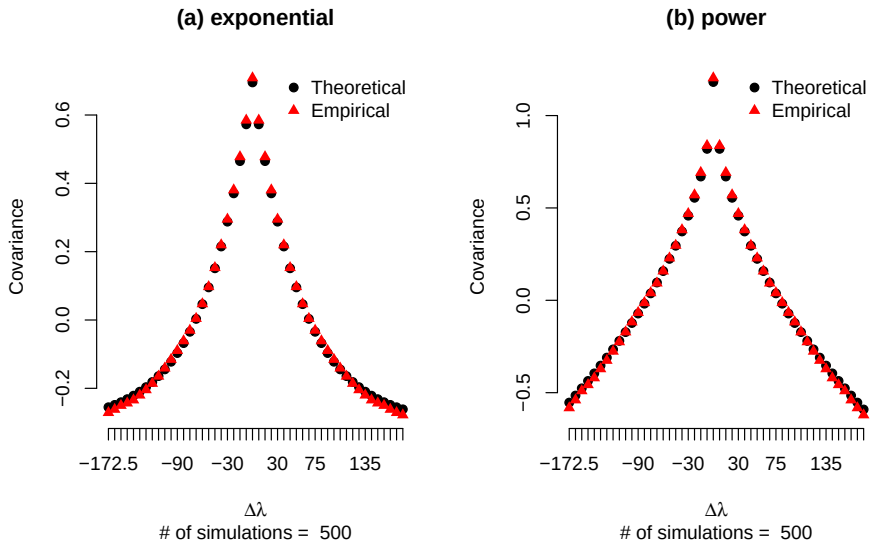


Figure 6. Theoretical and empirical covariance comparison on a circle using the modified covariance function $D(\theta)$.

be estimated by subtracting $\hat{a}_0$ form the original MOM estimator as given below.

$$\tilde{C}(\Delta\lambda) = \hat{C}(\Delta\lambda) - \hat{a}_0.$$

Our simulation shows both curves match very well (Figure 7).

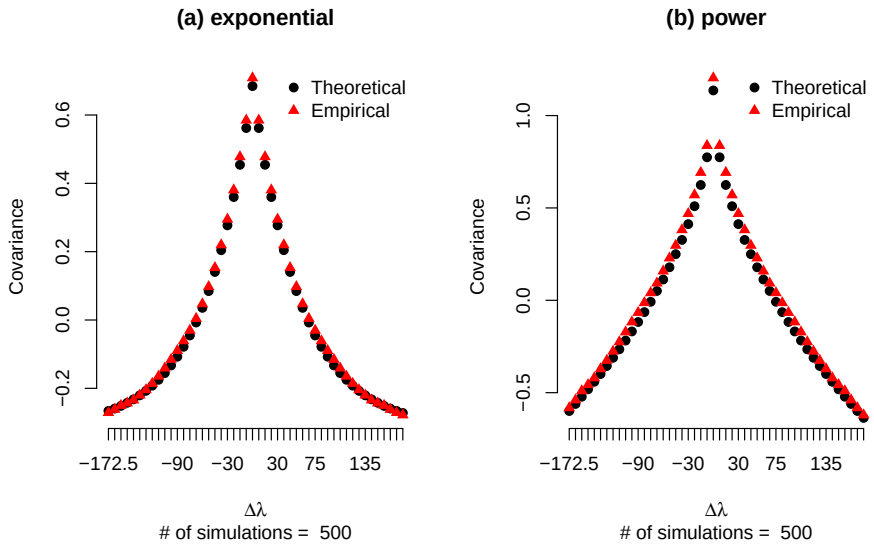


Figure 7. Theoretical and empirical covariance comparison on a circle after substrating $\hat{a}_0$ from the empirical covariance.

Now we theoretically calculate the unbiasedness of $\hat{C}(\Delta\lambda)$.

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

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 it's still equal to a_0 even if a0=0. In other hand, if $a_0 = 0$, the MOM estimator $\hat{C}(\Delta\lambda)$ is an asymptotically unbiased estimator of $C(\theta)$. This leads to the following proposition

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

3.4 Variogram MOM Estimator

When the random process on a circle is stationary, the semivariogram can be obtained by

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

In $\mathbb{R}^n$, The variogram MOM estimator generally performs better than the covariance MOM estimator (Cressie, 1993). Given gridded data observations $\underset{\sim}{X}$, the variogram MOM estimator is given by

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

We first perform a simulation with the same set up as before. Note that the theoretical exponential and power variogram functions are given below:

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

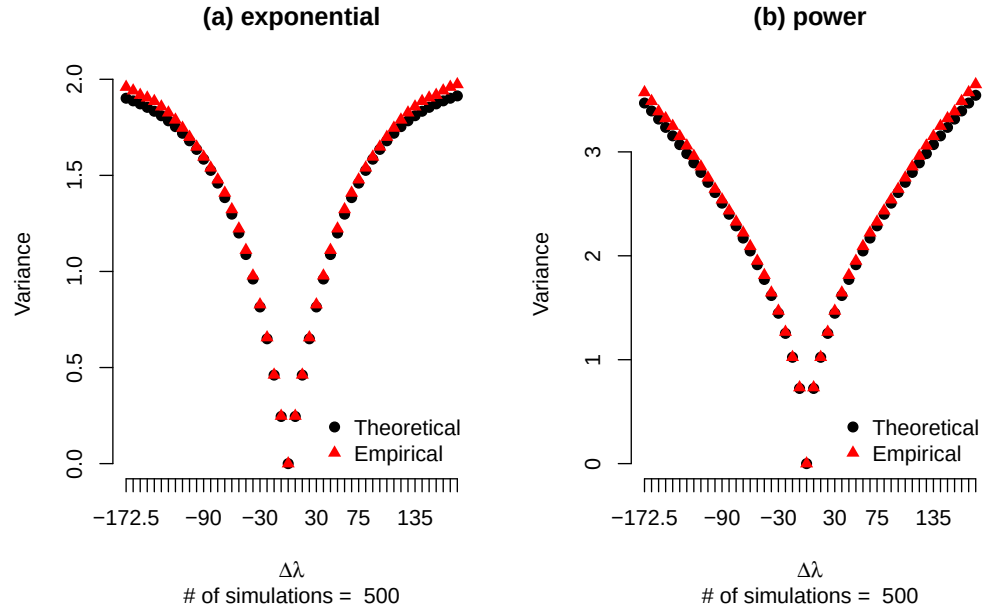


Figure 8. Comparison between theoretical and empirical variogram on a circle.

We compute the variogram estimator $\hat{\gamma}(\theta)$ with $n = 48$ gridded observations on the circle with 500 repetitions and then compared them with the theoretical values. The empirical and theoretical values match very well (Figure 8).

Now we explore the asymptotics of the MOM variogram estimator.

$$\begin{aligned} E(\hat{\gamma}(\Delta\lambda)) &= \frac{1}{2n}\sum_{i=1}^{n} E(X(t_i+\Delta\lambda) - X(t_i))^2 \\ &= \frac{1}{2n}\sum_{i=1}^{n} (2\gamma(\Delta\lambda)) = \gamma(\Delta\lambda). \end{aligned} \tag{3.5}$$

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 vectors. Assume that the random process $X(t)$ is stationary. $\hat{\gamma}(\Delta\lambda)$ can be written as

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

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

Let $n = 6$. For each of the four distance angles $\Delta\lambda = 0, \pi/3, 2\pi/3, \pi$, design matrices $A(\Delta\lambda)$ are given below:

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

similarly,

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

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

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

It follows that $A(\Delta\lambda) = A(m\delta)$ is a symmetric circulant matrix. From Section 1.3, the eigenvalues of $A(m\delta)$ are 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 is the trace of product of diagonal matrices, which is the sum of the product of corresponding eigenvalues from those circulant matrices.

Next 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,

$$\hat{\gamma}(\Delta\lambda) = (\underset{\sim}{X}-\underset{\sim}{\mathbf{1}}_n\mu)^T A(\Delta\lambda)(\underset{\sim}{X}-\underset{\sim}{\mathbf{1}}_n\mu). \tag{3.6}$$

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 Fourier matrix (orthonormal), solely depending on the dimension of A, and

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

If $\underset{\sim}{X}$ follows a multivariate normal distribution $N(\underset{\sim}{\mathbf{1}}_n\mu,\Sigma)$, 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:

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

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

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

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

That is, $\underset{\sim}{Y} \stackrel{\triangle}{=} (Y_1, Y_2, \cdots, Y_n)^T$ are independent normal random variates with mean 0 and variance $\lambda_j^{(\Sigma)}$, $j = 1, 2, \ldots, n$, respectively.

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 that $\frac{Y_m}{\sqrt{\lambda_m^{(\Sigma)}}} \sim N(0,1)$, 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

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

Here $\chi^2_{1,1}, \chi^2_{1,2}, \cdots, \chi^2_{1,n}$ are *i.i.d.* χ^2_1 random variables. Hence

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

that

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

which recovers the result(3.5) we obtained earlier. 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

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

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 $n \to \infty$. It should be pointed out that when $n \to \infty$, we are sampling denser and denser data points over the circle so that we have $\Delta\lambda$ obtainable. 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,

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

We summary our findings as the following proposition.

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

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

$$\begin{aligned} C_{n,m} &= \frac{1}{n}(1-\cos((m-1)\Delta\lambda)) \\ &\quad \left(C(0)+2\sum_{k=1}^{N-1} C(k\delta)\cos((m-1)k\delta)+C(\pi)\cos((m-1)\pi)\right)/(C(0)-C(\Delta\lambda)) \\ &\to (1-\cos((m-1)\Delta\lambda))\left(\frac{1}{\pi}\int_0^{\pi} C(\theta)\cos((m-1)\theta)d\theta\right)/(C(0)-C(\Delta\lambda)) \\ &= a_{m-1}2(1-\cos((m-1)\Delta\lambda))/(C(0)-C(\Delta\lambda)), \quad \text{as } n\to\infty. \end{aligned}$$

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 $n \geq N$,

$$\begin{aligned}
& P\left(\left|\sum_{m=1}^{n} C_{n,m}\chi^2_{1,m} - 1\right| \geq \varepsilon\right) \\
= \; & P\left(\sum_{m=1}^{n} C_{n,m}\chi^2_{1,m} > 1 + \varepsilon \quad \text{or} \quad \sum_{m=1}^{n} C_{n,m}\chi^2_{1,m} < 1 - \varepsilon\right) \\
\geq \; & P\left(\sum_{m=1}^{n} C_{n,m}\chi^2_{1,m} > 1 + \varepsilon\right) \geq P\left((b_2 - \varepsilon_0)\chi^2_1 > 1 + \varepsilon\right) \\
= \; & P\left(\chi^2_1 > \frac{1+\varepsilon}{b_2 - \varepsilon_0}\right) \not\to 0,
\end{aligned}$$

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

□

CHAPTER IV

PARAMETRIC MODELS AND ESTIMATION ON A SPHERE

4.1 Random Process on a Sphere

Suppose the process $\{X(P) : P \in \mathbb{S}^2\}$ ($\mathbb{S}^2$ unit sphere), defined in a common probability space, where $P = (\lambda, \phi) \in \mathbb{S}^2$ with longitude $\lambda \in [-\pi, \pi)$ and latitude $\phi \in [0, \pi]$, is continuous in quadratic mean with respect to the location P and has finite second moment. Then it can be represented by spherical harmonics (Huang et al., 2012; Jones, 1963; Li and North, 1997), with the sum converging in 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

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

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) (Leadbetter, 1967, page 83).

4.1.1 Homogeneous Covariance Functions on the Sphere

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

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

where θ_{PQ} is the spherical angle between two locations P, Q, 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 $C(\cdot)$ on the sphere must be non-negative definite, *i.e.,*

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

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

According to Schoenberg (1942), a real continuous function $C(\theta)$ is a valid homogeneous covariance function on the sphere ($\mathbb{S}^2$) if and only if it can be written in the

following form:

$$C(\theta) = \sum_{k=0}^{\infty} c_k P_k(\cos\theta), \quad \theta \in [0, \pi],$$

where $P_k(\cdot)$ is the Legendre polynomial, $\forall c_k \geq 0$ and $\sum_k c_k < \infty$. A general result of the above representation on $\mathbb{S}^d$ $(d > 2)$ can also be found in Schoenberg (1942).

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

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

Hence the coefficients c_k can be obtained as

$$c_k = \frac{2k+1}{2} \int_0^{\pi} C(\theta) P_k(\cos\theta) d\theta. \quad k = 0, 1, 2, \ldots \tag{4.1}$$

One can directly use the above integral to evaluate the validity of a homogenous covariance function on the sphere by checking if c_k is non-negative for all k 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 $(\mathbb{S}^2)$. For example, Huang et al. (2011) evaluated the validity of commonly used covariance models that are valid on $\mathbb{R}^d$ and summarized their findings in Table 2.

Table 2. Validity of covariance functions on the sphere, $a > 0, \theta \in [0, \pi]$

Model	Covariance function	Valid on $\mathbb{S}^2$
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

*When $\alpha \in (0, 1]$, the power model is valid on the sphere for some $c_0 \geq \int_0^{\pi} (\theta/a)^{\alpha} \sin\theta d\theta$.

Furthermore, Gneiting (2013) showed that the Matérn covariance function is only valid on the sphere when the smoothness parameter $\nu \in (0, 1/2]$. Another way of constructing a valid homogeneous covariance function on the sphere is by using the

valid covariance function in $\mathbb{R}^3$. Specifically, Yadrenko (1983) showed that if $K(\cdot)$ is a 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.

4.1.2 Variogram on a Sphere

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

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

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 $P_1, P_2, \ldots P_N \in \mathbb{S}^2$. Immediately from (4.1), for a continuous function $2\gamma(\cdot)$ with $\gamma(0) = 0$, the variogram is negative definite if and only if

$$\gamma(\theta) = \sum_{k=0}^{\infty} c_k(1 - P_k(\cos\theta)), \quad \theta \in [0, \pi] \tag{4.2}$$

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

It is 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 $\mathbb{S}^2$ Yaglom (1961) 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 the 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 the Method of Moments.

4.2 Axial Symmetry

For an axially symmetric process $X(P), P \in \mathbb{S}^2$ on the sphere, the covariance function $R(P, Q)$ at two locations $P = (\phi_P, \lambda_P), Q = (\phi_Q, \lambda_Q) \in \mathbb{S}^2$ 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 by Huang et al. (2012); Stein (2007), the covariance function can be expressed as the following:

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

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

$$R(P,Q) = R(\phi_P, \phi_Q, \Delta\lambda) = \sum_{m=-\infty}^{\infty} e^{im\Delta\lambda} C_m(\phi_P, \phi_Q), \quad m = 0, \pm 1, \pm 2, ..., \tag{4.4}$$

where $\Delta\lambda = \lambda_P - \lambda_Q \in [-\pi, \pi]$ and $\phi_P, \phi_Q \in [0, \pi]$. Here the complex bivariate continuous function $C_m(\phi_P, \phi_Q)$ has the following properties:

- Hermitian and positive definite.
- $\sum_{m=-\infty}^{\infty} |C_m(\phi_P, \phi_Q)| < \infty$ for 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 $R(P,Q)$,

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

Let $C_m(\phi_P, \phi_Q) = C_m^R(\phi_P, \phi_Q) + iC_m^I(\phi_P, \phi_Q)$. From Huang et al. (2012), 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:

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

4.2.1 Longitudinally Reversible Processes

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

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

we then call the underlying process to be longitudinally reversible (Stein (2007)). Now the covariance function $R(P, Q)$ of a real-valued longitudinally reversible process reduces to

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

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 Section 4.2, many covariance models valid in $\mathbb{R}^d$ might not be valid in $\mathbb{S}^2$. Therefore, it is necessary to develop parametric models that possibly reflect the topological structure of compactness of a sphere. Although the Matérn covariance model has been often used to model global data in recent years, it has some drawbacks. For example, it has been shown that the homogeneous Matérn covariance model is not valid when the smoothness parameter is larger than 0.5. Further modifications have been proposed e.g., (Jeong and Jun, 2015; Jun and Stein, 2008; Li, 2013. Huang et al. (2012) discussed a new representation for the covariance structure of an axially symmetric process. 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 (4.4), is clearly a function of both latitudes and the difference of longitudes. If we assume that in (4.4) $C_m(\phi_P,\phi_Q) = a_m\tilde{C}(\phi_P-\phi_Q)$ only depends on the difference between ϕ_P and ϕ_Q, such that $C_1(\Delta\lambda) = \sum_{m=-\infty}^{+\infty} a_m e^{im\Delta\lambda}$ exists, then $R(P,Q) = C_1(\Delta\lambda)\tilde{C}(\phi_P-\phi_Q)$. This is a simple separable model as given in Huang et al. (2011). Here is an example when both covariance components are exponential

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

where $c_0 > 0$, and both $a > 0$ and $b > 0$ are defined as decay parameters in longitude and latitude, respectively. However, the separable models are not capable to capture the covariance structure of the entire sphere since it obvious also assumes the stationarity across latitudes. Huang et al. (2012) further proposed some non-separable covariance models by carefully choosing functions for $C_m(\phi_P,\phi_Q)$ that are valid on the sphere, under which some $R(P,Q)$ are given below:

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

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

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

where $A > 0, 0 < p < 1, u \in R$ are all constants, $\Theta = \Delta\lambda + u(\phi_P-\phi_Q) - 2k\pi$, and k is chosen such that $\Theta \in [0,2\pi]$. First notice that all of the three covariance models (4.6, 4.7, 4.8) depend not only on the difference in longitudes, but on the difference in latitudes as well. As an example, we consider the model (4.6) from Huang et al.

(2012). When $\phi_P = \phi_Q$, the model reduces to

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

indicating all covariance functions on latitudes are the same. If we further set $\Delta\lambda = 0$, the variance of the process over all locations is given by,

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

implying a constant variance across the entire sphere. This is unrealistic, since we have seen from MSU and TOMS data that the variance highly depends on latitudes. As a generalization of the above models, we propose the following proposition.

Proposition 4.1. *Let $C(\cdot) = C(x-y)$ be the stationary covariance function on $\mathbb{R}^d$. For simplicity, we let $f(\omega) \geq 0$ be the spectral density of $C(\cdot)$. Then*

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

$$\textit{with } C_2 \geq \int_{-\infty}^{\infty} f(\omega)d\omega > 0,$$

is the non-stationary covariance function on $\mathbb{R}^d$.

Proof. by Bochner's theorem, if $f(\cdot) \geq 0$ is the spectral density of the covariance function $C(\cdot)$, we have

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

Now for any N, 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 $C_2 = \int_{-\infty}^{\infty} f(\omega)d\omega$,

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

This proves the positive definiteness of $\tilde{C}(\cdot,\cdot)$, which concludes the proof.

□

We apply the above proposition on the following two stationary covariance functions,

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

where $C > 0, a > 0$. We arrive with the following two non-stationary covariance functions.

$$\tilde{C}(\phi_P, \phi_Q) = C_1(C_2 - e^{-a|\phi_P|} - e^{-a|\phi_Q|} + e^{-a|\phi_P - \phi_Q|}), \tag{4.9}$$

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

where $C_1, a > 0$, 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:

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

Therefore, we propose the following covariance models for axially symmetric processes on the sphere, which generalize the models given by (Huang et al., 2012, model 1, model 4, model 5):

$$R(P,Q) = \tilde{C}(\phi_P, \phi_Q)\frac{1-p^2}{1-2p\cos\Theta + p^2}, \tag{4.11}$$

$$R(P,Q) = \tilde{C}(\phi_P, \phi_Q)\log\frac{1}{(1-2p\cos\Theta * + p^2)}, \tag{4.12}$$

$$R(P,Q) = \tilde{C}(\phi_P, \phi_Q)\left(\frac{\pi^4}{90} - \frac{\pi^2\Theta^2}{12} + \frac{\pi\Theta^3}{12} - \frac{\Theta^4}{48}\right), \tag{4.13}$$

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

Remark 6. 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 $\lambda_{min} = \pi + \lambda_{max}$.

Remark 7. The scaling parameter p changes rapidly at the supremum and infimum of the covariance models and parameters C_1, C_2, and a are regular scaling parameters. The parameter u is a location parameter which shifts the covariance from left to right

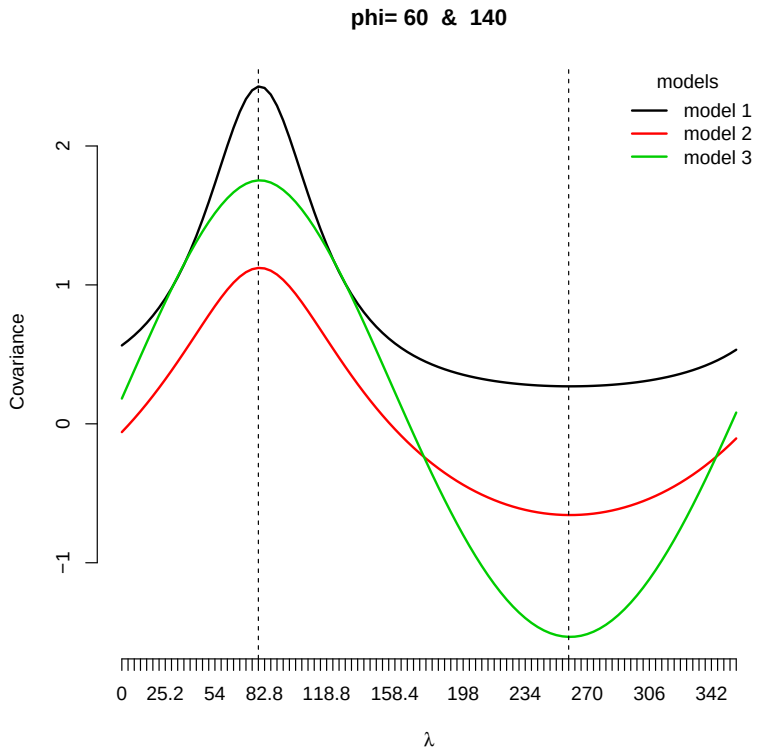


Figure 9. The covariance between 30^0S and 50^0N (latitudes 60^0 and 140^0) of three covariance models with exponential family, *i.e.*, $\tilde{C}(\phi_P, \phi_Q)$ given by (4.9) over 100 longitudes (we set $C_1 = C_2 = a = u = 1$, and $p = 0.5$).

with respect to $\Delta\lambda$) when $u > 0$. When $u = 0$ it provides a longitudinally reversible covariance model.

4.4 Covariance and Variogram Estimators on a Sphere

The covariance function $R(P, Q)$ has also be termed as a cross covariance in the literature, as it captures the covariance of the process between points at two latitudes with longitudes separated by $\Delta\lambda \in (-\pi, \pi)$. In this section, we consider the estimator for the covariance function $R(P, Q)$ based on MOM as well as its properties. In addi-

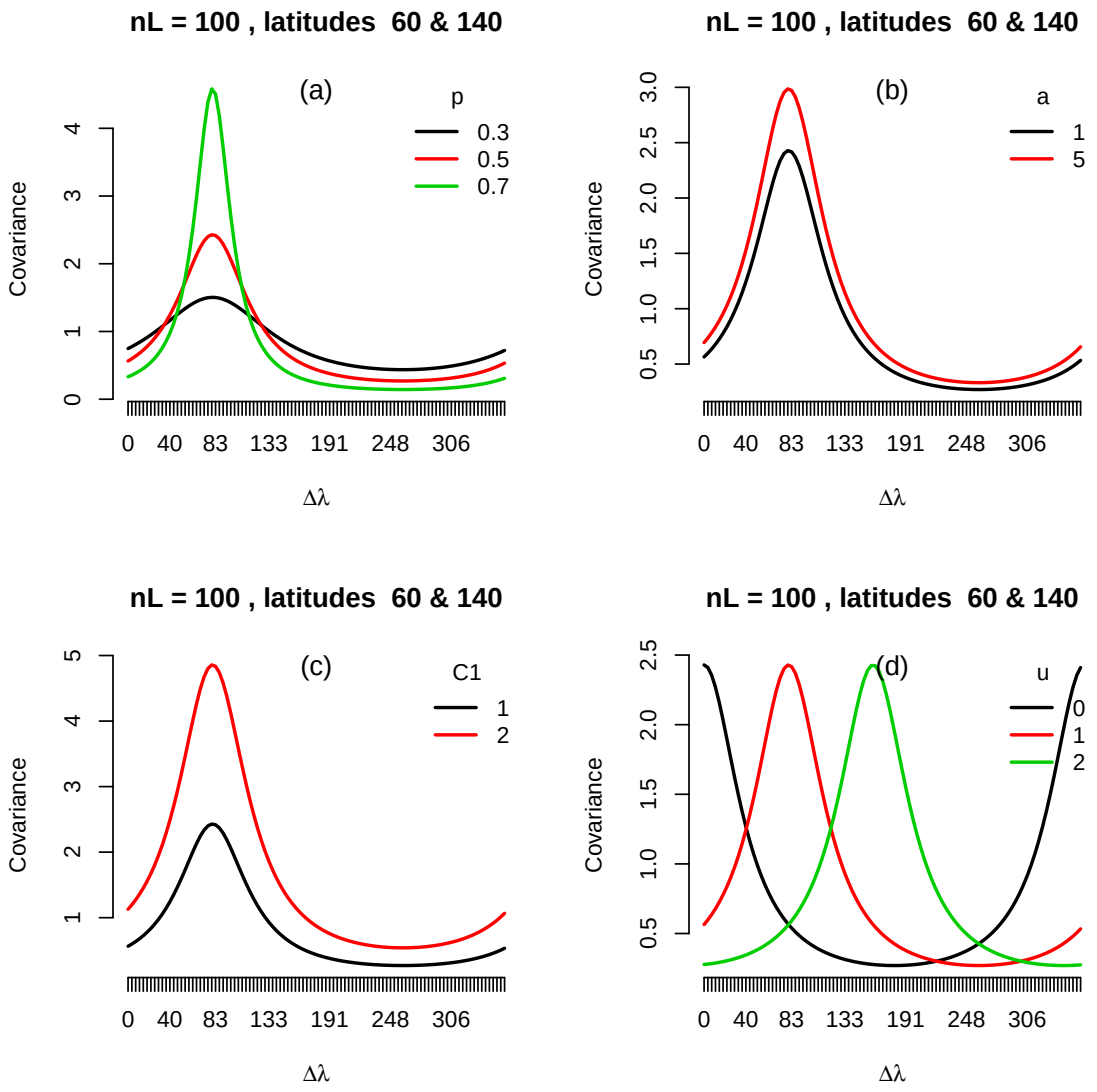


Figure 10. Covariance curves for different parameters using model 1 (Huang et al., 2012): (a)-parameter p, (b)-parameter a, (c)-parameter C_1 (similar pattern for parameter C_2), (d)-parameter u

tion, we will introduce the cross variogram function on the sphere. The unbiasedness of the cross variogram estimator based on MOM will also be discussed.

4.4.1 Cross Covariance and Cross Variogram Functions

As discussed in Wackernagel (2013), the cross covariance function, or $R(P, Q)$, can be decomposed as the following:

$$\begin{aligned} R(\phi_P, \phi_Q, \Delta\lambda) &= \frac{1}{2}(R(\phi_P, \phi_Q, \Delta\lambda) + R(\phi_P, \phi_Q, -\Delta\lambda)) \\ &\quad + \frac{1}{2}(R(\phi_P, \phi_Q, \Delta\lambda) - R(\phi_P, \phi_Q, -\Delta\lambda)). \end{aligned}$$

For fixed latitudes ϕ_P and ϕ_Q, the first term is an even function of $\Delta\lambda$, while the second term is an odd function of $\Delta\lambda$. We notice that, when $\phi_P \neq \phi_Q$, $R(\phi_P, \phi_Q, \Delta\lambda)$ has the following properties:

- $R(\phi_P, \phi_Q, \Delta\lambda) \neq R(\phi_Q, \phi_P, \Delta\lambda)$;
- $R(\phi_P, \phi_Q, -\Delta\lambda) \neq R(\phi_P, \phi_Q, \Delta\lambda)$;
- $R(\phi_P, \phi_Q, -\Delta\lambda) = R(\phi_Q, \phi_P, \Delta\lambda)$.

As a special case, when $\phi_P = \phi_Q$ fixed, two locations are on the same circle, the cross covariance $R(\phi_P, \phi_P, \Delta\lambda)$ is a stationary covariance function on the circle, a function depending on longitudinal difference ($\Delta\lambda$) only.

Now we introduce the cross variogram function on the sphere. The cross variogram function is defined as following:

$$2\gamma(\phi_p, \phi_Q, \Delta\lambda) = E\left((X(\phi_P, \lambda + \Delta\lambda) - X(\phi_P, \lambda))(X(\phi_Q, \lambda + \Delta\lambda) - X(\phi_Q, \lambda))\right).$$

When $\phi_P = \phi_Q$, the above expression reduces to the expression for the variogram function on the circle.

Note that

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

That is,

$$\gamma(\phi_p, \phi_Q, \Delta\lambda) = R(\phi_p, \phi_Q, 0) - \frac{1}{2}(R(\phi_P, \phi_Q, \Delta\lambda) + R(\phi_P, \phi_Q, -\Delta\lambda)). \tag{4.14}$$

We notice that the (semi-)cross variogram relates to only the even term of the cross covariance function. Wackernagel (2013) argues that cross variogram might not be sufficient when there is a delayed affect, which introduces the non-zero value representing from the odd component in the cross covariance decomposition.

When the axially symmetric process $X(P)$ is also longitudinally reversible, that is, $R(\phi_P, \phi_Q, -\Delta\lambda) = R(\phi_P, \phi_Q, \Delta\lambda)$, we have

$$\gamma(\phi_p, \phi_Q, \Delta\lambda) = R(\phi_p, \phi_Q, 0) - R(\phi_P, \phi_Q, \Delta\lambda),$$

which is the same relationship as the one between the covariance and variogram functions on the circle.

4.4.2 The MOM Cross Covariance and Cross Variogram Estimators

We now provide the MOM estimators for the cross covariance $R(\phi_P, \phi_Q, \Delta\lambda)$ and cross semivariogram function $\gamma(\phi_P, \phi_Q, \Delta\lambda)$ of an axially symmetric process on the sphere. First for any two latitudes ϕ_P and ϕ_Q with $\{\lambda_i, i = 1, 2, \ldots, n\}$ representing the gridded longitudes on each circle (for simplicity, we assume $n = 2N$ is an even number), then the MOM estimator $\hat{R}(\phi_P, \phi_Q, \Delta\lambda)$ is given by

$$\hat{R}(\phi_P, \phi_Q, \Delta\lambda) = \frac{1}{n}\sum_{i=1}^{n}(X(\phi_P, \lambda_i + \Delta\lambda) - \bar{X}_P)(X(\phi_Q, \lambda_i) - \bar{X}_Q), \tag{4.15}$$

where $\Delta\lambda = 0, 2\pi/n, 4\pi/n, \cdots, 2(N-1)\pi/n$, $\bar{X}_P = \frac{1}{n}\sum_{i=1}^{n} X(\phi_P, \lambda_i)$, and a similar expression for $\bar{X}_Q$.

Now we show the unbiasedness of the cross covariance estimator.

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

where $\Lambda_i = \lambda_i + \Delta\lambda$ and note that,

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

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

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

Note 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,

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

We summarize the above calculations as the following proposition.

Proposition 4.2. *The MOM cross covariance estimator is biased with the constant shift given by $C_0(\phi_P,\phi_Q)=cov(\bar{X}_P,\bar{X}_Q)$. Hence, the true $R(P,Q)$ may not be identifiable based on the MOM estimator.*

Remark 8. When $\phi_P = \phi_Q$, the result above reduces to that for the MOM covariance estimator on the circle.

Remark 9. If the mean at each latitude is zero, then we can rewrite the cross covariance MOM estimator as following:

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

One can show that this estimator is unbiased.

Next we consider the MOM cross semivariogram estimator for axially symmetric processes on the sphere.

$$\hat{\gamma}(\phi_p, \phi_Q, \Delta\lambda) = \frac{1}{2n}\sum_{i=1}^{n} \left(X(\phi_P, \Lambda_i) - X(\phi_P, \lambda_i))(X(\phi_Q, \Lambda_i) - X(\phi_Q, \lambda_i))\right), \quad (4.16)$$

where $\Lambda_i = \lambda_i + \Delta\lambda$. We have

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

which is unbiased. Here we summarize our finding as follow.

Proposition 4.3. *The MOM cross semivariogram estimator is unbiased.*

Remark 10. We expect the similar result as the one in the case of ciccle for the consistency of the MOM cross semivariogram estimator on the sphere, but the proof will be more complicated. This will be one of the areas for our future research.

CHAPTER V

GLOBAL DATA GENERATION

5.1 Introduction

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 data generation, there seems only a limited research in literature. In order to capture non-stationarity, Jun and Stein (2007) proposed spatio-tempo covariance functions on the sphere by applying first order differential operator to fully symmetric spatio-tempo processes on sphere. Further, Jun and Stein (2008) extended the above approach and used the Discrete Fourier Transform (DFT) to find out the inverse of the covariance matrix when calculating the exact likelihood for the data on regular grids. They indicated that the inverse of the covariance matrix for axially symmetric data is of the order of $O(n_l^3 n_L)$ where n_L is the number of longitudes and n_l the number of latitudes. However, no data generation and simulation was discussed. Li (2013) proposed convolution methods to generate random fields with a class of Matérn-type kernel functions by allowing the parameters in the kernel function to vary with latitudes. They conducted a simulation study by generating data based on their proposed method. They used maximum likelihood estimation and local smoothing to recover the given parameters. In the recent work by Jeong and Jun (2015), they presented a number of simulation scenarios using the Matérn-like covariance models that include stationary and non-stationary processes.

They seemed to generate data for simulations based in the given covariance structure directly. No data validation was given. As we have discussed in the previous chapters, the global data normally exhibit both complexity and high dimensionality. For example, the MSU data was observed on a 72×144 grid which results in an estimated covariance matrix with a dimension of 10368×10368. Hence, it is necessary to develop an algorithm that is efficient and with low dimensionality. In this research, we use the Discrete Fourier Transform to decompose the process as Fourier series on the circles, and represent the random Fourier coefficients with circularly-symmetric complex random vectors. This has greatly reduced both computational cost as well as dimensionality. It also gives a better insight about the axially symmetric process on the sphere.

This chapter is organized as follow. We first layout the details and methodologies of generating data on a sphere using circularly-symmetric random vectors. Then we provide an algorithm for global data generation process. Finally, we conduct simulations and use the cross variogram and cross covariance MOM estimation to validate the data generated.

5.2 Method Development

Let $X(P)$ be a continuous real-valued Gaussian random process defined on a unit sphere $\mathbb{S}^2$, where $P = (\lambda, \phi) \in \mathbb{S}^2$ with longitude $\lambda \in [-\pi, \pi)$ and latitude $\phi \in [0, \pi]$. Following Remark 2.5 in Huang et al. (2012), for each fixed latitude ϕ, $X(P)$ can then be represented as a stationary process on the circle. More specifically,

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

where

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

with $\mathrm{E}(W_m(\phi_P)\overline{W_n(\phi_Q)}) = \delta(m,n)C_m(\phi_P,\phi_Q)$. That is, $\{W_m(\phi), \phi \in [0,\pi], m = 0, \pm1, \pm2, \ldots,\}$ are uncorrelated Gaussian complex random variables with covariance function given by $C_m(\phi_P,\phi_Q)$.

In order to obtain axially symmetric random variates for a given latitude ϕ, we first construct normal independent (complex) random variates $W_m(\phi)$ that follow the given variance-covariance function $C_m(\phi_P,\phi_Q)$. To achieve this, we truncate the infinite summation in (5.1) to obtain a finite summation up to N as given below (with the abuse of notation $X(P)$ for notation simplicity).

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

Since W_m's are independent for $m = 0, \pm1, \pm2, \cdots, \pm N$, we have

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

indicating that the Fourier series for covariance function $R(P,Q)$ has also been truncated up to N. Let $W_m(\phi) = W_m^r(\phi) + iW_m^i(\phi)$ and $C_m(\phi_P,\phi_Q) = C_m^r(\phi_P,\phi_Q) + iC_m^i(\phi_P,\phi_Q)$.

In order to obtain real data values that follow the given $R(P,Q)$, we require that

- $W_{-m}(\phi) = \overline{W_m(\phi)}$, and
- $C_{-m}(\phi_P, \phi_Q) = \overline{C_m(\phi_P, \phi_Q)}$, for $m = 0, 1, 2, \cdots, N$.

First we simplify the process.

$$\begin{aligned} X(P) &= \sum_{m=-N}^{N} W_m(\phi) e^{im\lambda} = W_0(\phi) + \sum_{m=1}^{N} W_m(\phi) e^{im\lambda} + \sum_{m=-1}^{-N} W_m(\phi) e^{im\lambda} \\ &= W_0(\phi) + \sum_{m=1}^{N} W_m(\phi) e^{im\lambda} + \sum_{m=1}^{N} \overline{W_m(\phi)} e^{-im\lambda} \\ &= W_0(\phi) + \sum_{m=1}^{N} \big[(W_m^r(\phi) + iW_m^i(\phi))(\cos(m\lambda) + i\sin(m\lambda)) \\ &\quad + (W_m^r(\phi) - iW_m^i(\phi))(\cos(m\lambda) - i\sin(m\lambda)) \big] \\ &= W_0(\phi) + 2\sum_{m=1}^{N} \big[W_m^r(\phi)\cos(m\lambda) - W_m^i(\phi)\sin(m\lambda) \big]. \end{aligned} \tag{5.3}$$

Now we consider the corresponding variance-covariance function for each $W_m(\phi)$.

Note that to have the real-valued data observations, hence

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

Writing $C_m(\phi_P, \phi_Q) = C_m^r(\phi_P, \phi_Q) + iC_m^i(\phi_P, \phi_Q)$ and with (5.4), 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,

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

In addition, we set the following.

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

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

The above relationships (5.5) and (5.6) will become our basis for data generation.

5.2.1 Data Generation

Now for each fixed $m = 0, 1, 2, \cdots, N$, we write $W_m(\phi) = W_m^r(\phi) + iW_m^i(\phi)$ then $\overline{W_m}(\phi) = W_m^r(\phi) - iW_m^i(\phi)$. 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 $\sigma_m^2(\phi) = \frac{1}{2}C_m^r(\phi, \phi)$ and $C_m^i(\phi, \phi) = 0$ (that is, $W_m(\phi)$ is circularly symmetric). 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

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

We now show that the Gaussian random vector $\underset{\sim}{W}_m$ is circularly-symmetric, that is, 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 of $\underset{\sim}{W}_m$) and pseudo-covariance $M_W = E(\underset{\sim}{W}_m \underset{\sim}{W}_m^T)$, and show that $M_W = 0$. First note that

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

We calculate each of the above entries. For any i, j,

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

The last equality is due to the independence of $W_m^r(\phi_i)$ and $W_m^i(\phi_i)$. Therefore, $M_W = 0$, and so $\underbrace{W}_{\sim m}$ is circularly-symmetric. In addition,

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

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

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

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

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

We arrive

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

Note that K_V is a non-negative definite matrix, its singular decomposition is given as follows.

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

where Λ is a diagonal matrix with eigenvalues (real-positive) of K_V and P is an orthonormal matrix, containing the corresponding orthonormalized eigenvectors of K_V. Let $A = K_V^{1/2} = P\Lambda^{1/2}P^T$ and let $\underset{\sim}{Z} = \{z_1^{(1)}, z_2^{(1)}, \ldots, z_{n_l}^{(1)}, z_1^{(2)}, z_2^{(2)}, \ldots, z_{n_l}^{(2)}\}$ be a vector of *i.i.d.* standard normal random variates, we obtain

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

and hence $\underset{\sim}{W}_m$.

Now for each latitude $\phi_l, l = 1, 2, \cdots, n_l$ and $\lambda_k, k = 1, 2, \cdots, n_L$, we denote the axially symmetric real data as $\{X(\phi_l, \lambda_k)\}$. These random variates can be obtained

from the equation (5.3), which is rewritten as below.

$$X(\phi_l, \lambda_k) \;=\; W_0(\phi_l) + 2\sum_{m=1}^{N} \left[W_m^r(\phi_l)\cos(m\lambda_k) - W_m^i(\phi_l)\sin(m\lambda_k)\right]. \quad (5.7)$$

Remark 11. For the above decomposition, one needs to compute the eigenvalues and eigenvectors of K_V, which has the computational cost of $O((n_l)^2)$.

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

- *Given a cross covariance function $R(P,Q)$ with given parameters (C_1, C_2, a, u, p)*
- *Choose a resolution $\phi_1, \ldots, \phi_{n_l}, \lambda_1, \ldots, \lambda_{n_L}$ (or $n_l \times n_L$)*
- *Obtain $\{W_m\}$,*
  - (1) *derive $C_m(\phi_P, \phi_Q)$ $m = 0, 1, \ldots, n_L/2$ based on $R(P,Q)$*
  - (2) *for each m get $Re(K_W)$ and $Im(K_W)$ hence obtain K_V*
  - (3) *use SVD to compute $\underset{\sim}{V}_m$ $(n_l - tuples)$*
  - (4) *get $\underset{\sim}{W}_m$'s from $\underset{\sim}{V}_m$*
- *Apply the equation (5.7) to obtain gridded data.*

5.3 Simulation Setup

We now implement the algorithm above to obtain axially symmetric data with the given covariance structure. We will use models (4.11), (4.12), and (4.13) discussed in Section 4.3 in the simulation. We select two sets of parameters, which are given below.

Table 3. Parameter values

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

The parameter u is the location parameter and $u = 0$ gives the covariance function of a longitudinally reversible process on a sphere.

5.3.1 Deriving $\mathbf{C_m}$ for Model 1

As we note that it is very crucial to derive $C_m(\phi_P, \phi_Q)$ based on the given $R(P, Q)$. Here is the detailed steps to derive $C_m(\phi_P, \phi_Q)$ from model 1 (4.11). $C_m(\phi_P, \phi_Q)$ for other models can be obtained similarly.

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

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 $m \neq 0$,

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

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 $\tilde{C}(\phi_P, \phi_Q)$ is given by (4.9) then $C_m(\phi_P, \phi_Q)$ for model 1 is given by,

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

The $C_m(\phi_P, \phi_Q)$ functions for all three proposed models are given below,

Table 4. The $C_m(\cdot, \cdot)$ functions for covariance models used in data generation

Model	$C_m(\phi_P, \phi_Q)$	Parameters
model 1	$\tilde{C}(\phi_P, \phi_Q) p^m$ and $C_0 = \tilde{C}(\phi_P, \phi_Q)$	$m = 0, \pm 1, \pm 2, ... \quad p \in (0, 1)$
model 2	$\tilde{C}(\phi_P, \phi_Q) \frac{p^m}{m}$ and $C_0 = 0$	$m = \pm 1, \pm 2, ... \quad p \in (0, 1)$
model 3	$\tilde{C}(\phi_P, \phi_Q) \frac{1}{m^4}$ and $C_0 = 0$	$m = \pm 1, \pm 2, ...$

Note that out of above three models, model 1 has $C_0(\phi_P, \phi_Q) \neq 0$. Hence, from Chapter 4, the MOM cross covariance estimator is biased with the non-estimable shift. On the other hand, both models 2 and 3 have $C_0 = 0$, and hence one could use the unbiased MOM cross covariance estimator for data validation.

5.3.2 Data Generation Through R(P, Q)

In order to demonstrate how effective our data generation algorithm is, we also consider the data generated through the given covariance matrix $R(P, Q)$ directly.

Note that the covariance matrix $R(P,Q)$ is a real block circulant matix with the following (Jun and Stein, 2008; Li, 2013).

$$R(P,Q) = \begin{bmatrix} R_0 & R_1 & R_2 & \cdots & R_{n_L-1} \\ R_{n_L-1} & R_0 & R_1 & \cdots & R_{n_L-2} \\ R_{n_L-2} & R_{n_L-1} & R_0 & \cdots & R_{n_L-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_1 & R_2 & R_3 & \cdots & R_0 \end{bmatrix}_{n_l n_L \times n_l n_L} \tag{5.8}$$

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 longitude 1, R_1 is the covariance matrix between latitudes at longitude 1 and longitude 2, and so on. Note that each R_j is symmetric only when the process is longitudinally reversible. To generate the gridded data values $\underset{\sim}{X}$ which is formed by $\{X(\phi_i, \lambda_j), i = 1, 2, \ldots, n_l, j = 1, 2, \ldots, n_L\}$, one can just simply obtain a vector $\underset{\sim}{Z}$ that contains *i.i.d.* standard normal variates of size $n_l \times n_L$, then

$$\underset{\sim}{X} = R^{1/2}(P,Q) * \underset{\sim}{Z}.$$

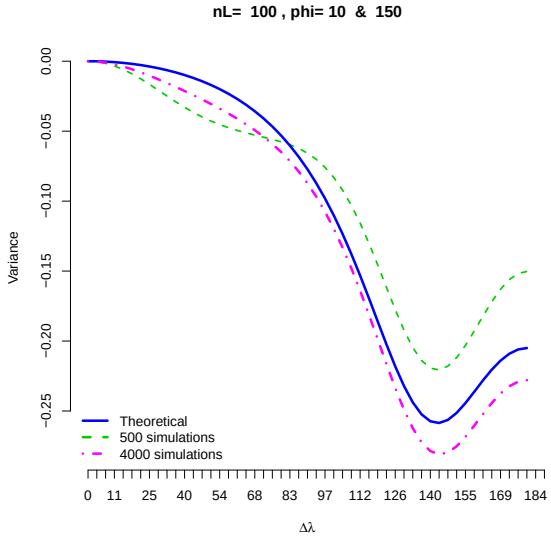
5.4 Results

To validate if the generated axially symmetric global data follow the given covariance model, we compare the MOM cross variogram estimates from the generated data to its theoretical counterparts. As indicated in Chapter 4, although the cross variogram (4.16) only reflects the even component of the cross covariance, it is unbiased. Moreover, when the data are generated from the longitudinally reversible covariance function, the cross variogram is equivalent to the cross covariance.

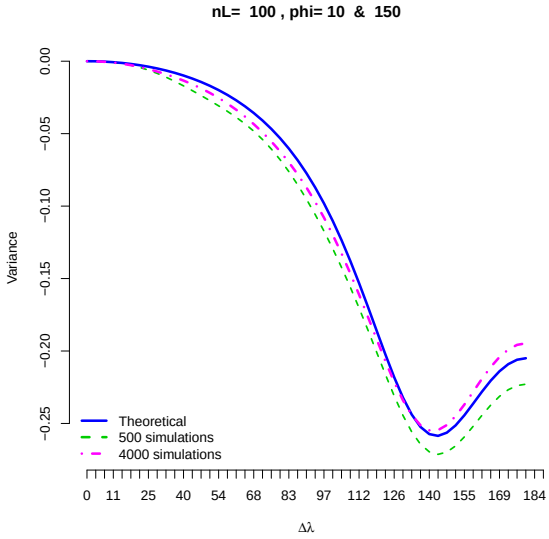
We have conducted the simulation for different pairs of latitudes with the same number of longitudes ($n_L = 100$). The pair of latitudes has been selected from very close by to far away. The cross variogram estimates are almost identical to the theoretical values when the pair of latitudes are closer, which is not shown in the dissertation. Here we demonstrate a case with larger latitude difference ($\phi_P = 10^0, \phi_Q = 150^0$ equivalent to 70^0S and 60^0N) in order to capture the largest possible errors.

5.4.1 Comparison of MOM Estimators

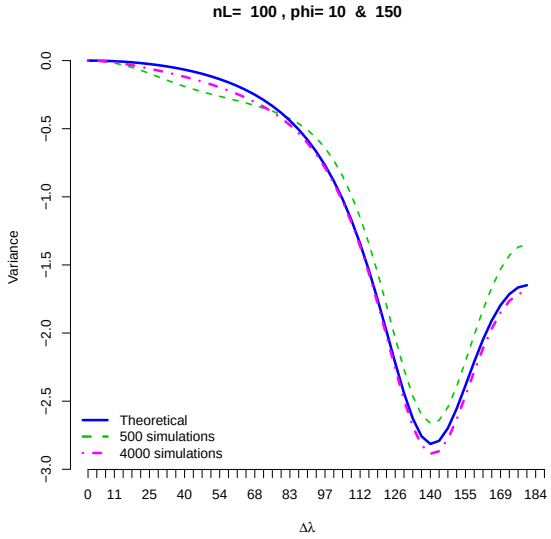
Now we compare the cross variogram estimates given in (4.16), using our approach (with C_m) and the approach with $R(P, Q)$ directly to generate data, respectively. One can see that the estimate given by our method is closer to the true values than those from $R(P, Q)$ directly, under 500 repetitions and 4000 repetitions respectively (Figure 11).



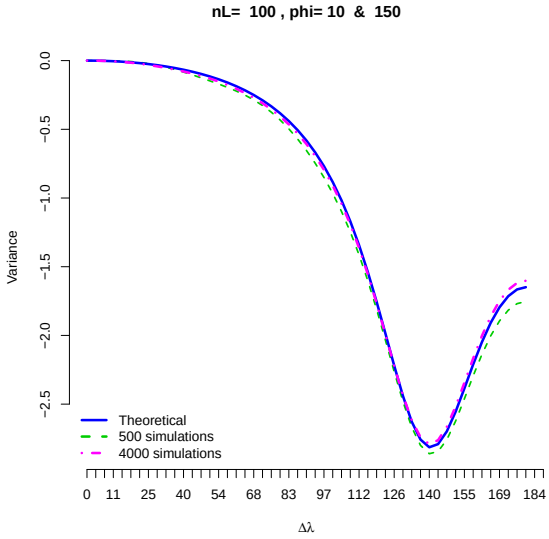
(a) Using paramter Set 1 and $R(P,Q)$



(b) Using paramter Set 1 and C_m



(c) Using paramter Set 2 and $R(P,Q)$



(d) Using paramter Set 2 and C_m

Figure 11. Using parameter Set 1 and Set 2 to perform the variogram estimator under model 1, solid line (blue) represents the theoretical values of cross variogram and dashed lines (green, purple) represents the estimates for 500 and 4000 simulations respectively.

5.4.2 Results for Longitudinally Reversible Processes

Setting parameter $u = 0$ in all of our models yields longitudinally reversible covariance functions (see Figure 10 1(d)). Based on 500 simulations, one can see that the cross variogram estimates from direct $R(P, Q)$ approach are slightly away from the theoretical values all three models (Figure 12). In contrast from the C_m approach the estimates are very close to theoretical values (Figure 13).

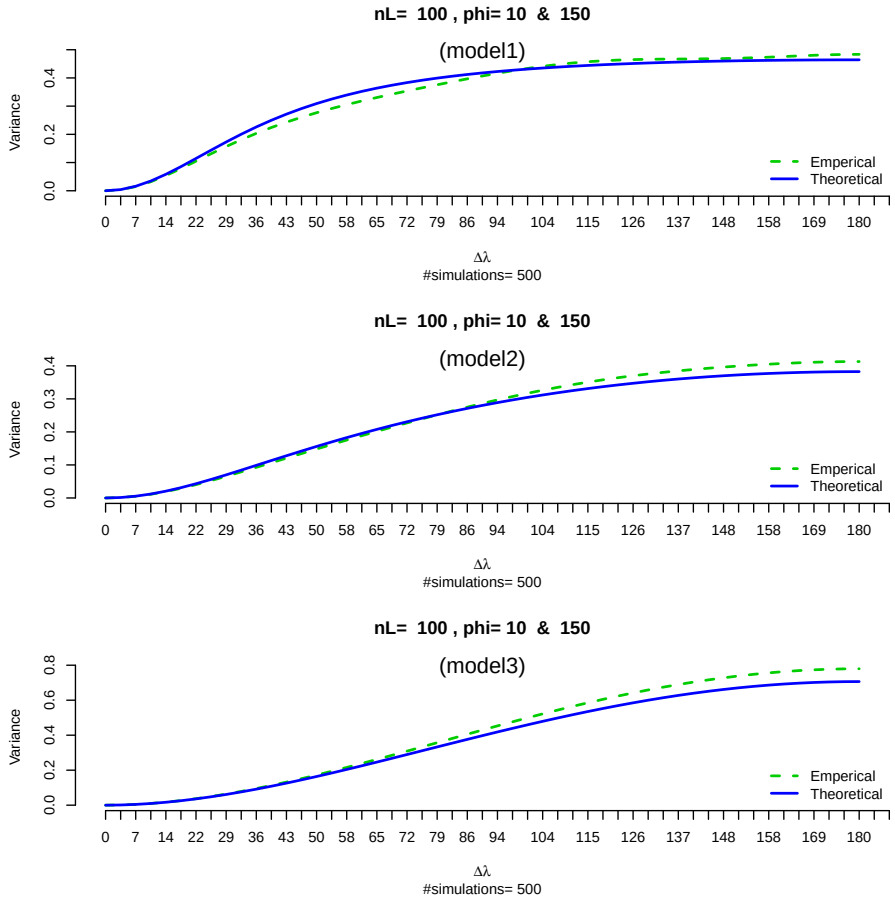


Figure 12. Based on $R(P, Q)$ approach the cross variogram estimator comparison for longitudinally reversible process using model 1, model 2, and model 3 (when $u = 0$).

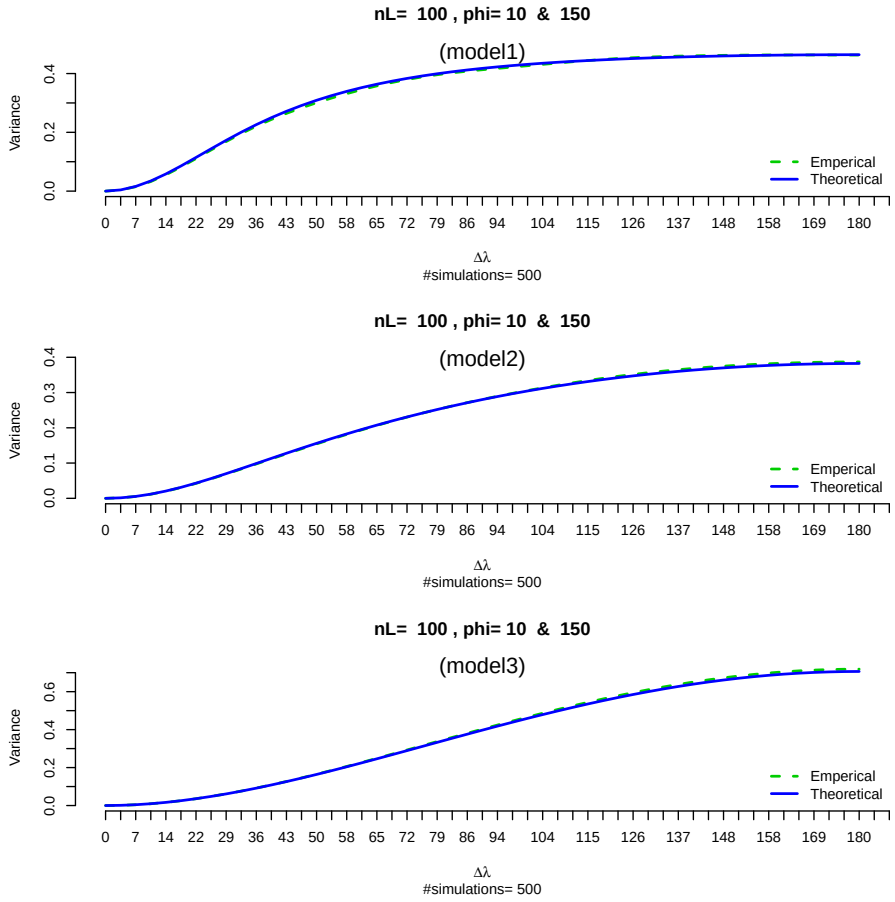


Figure 13. Based on C_m approach the cross variogram estimator comparison for longitudinally reversible process using model 1, model 2, and model 3 (when $u = 0$).

5.4.3 Comparison of Cross Covariance

As indicated from Chapter 4, when $C_0(\phi_P, \phi_Q) = 0$, the cross covariance MOM estimator is unbiased. Therefore we obtain the cross covariance MOM estimates given by (4.15) under model 2 and model 3, and then compare them with the true values. Here we select two pairs of latitudes, $\phi = 70, 80$ (20^0S ,10^0S) and $\phi = 60, 120$ (30^0S,

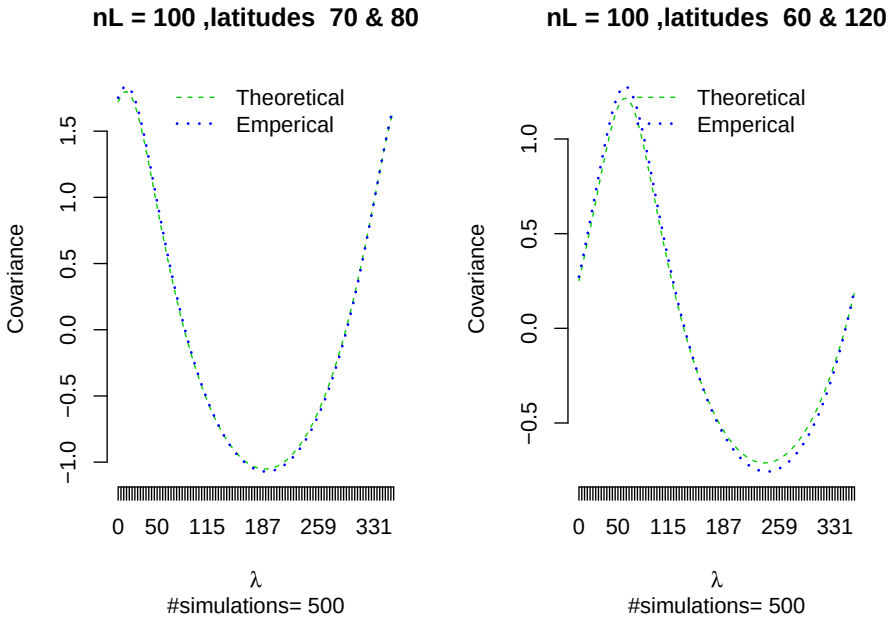
40^0N). One can note that the cross covariance estimates match with the theoretical values very well (Figure 14).

5.4.4 Comparison of MSE

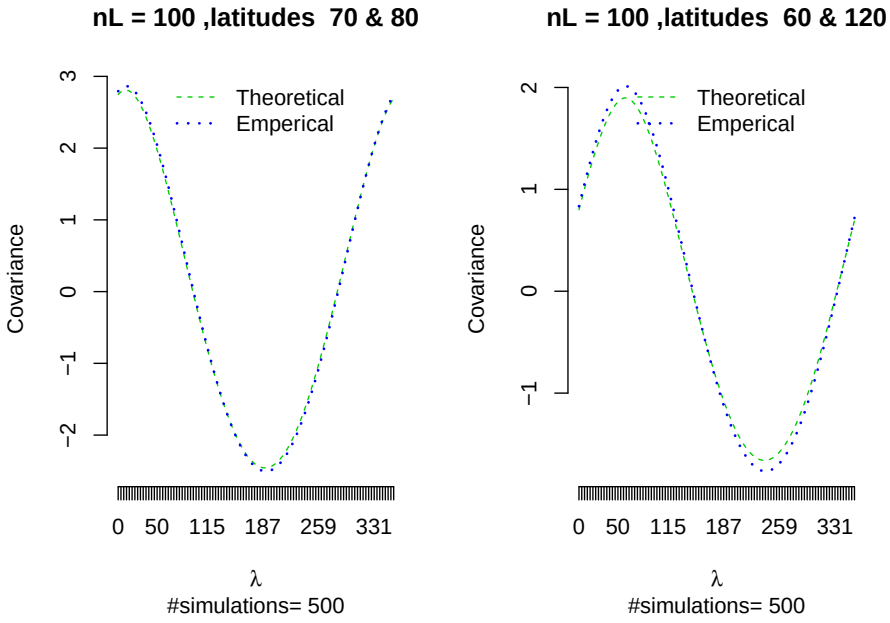
In addition to comparing the biases, we also consider the mean square error (MSE) of the MOM cross variogram estimates obtained under both C_m and direct $R(P,Q)$ approaches. The overall MSE is calculated based on the following formula.

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

where $\delta = \frac{2\pi}{n_L}$ and nn is the number of simulations.

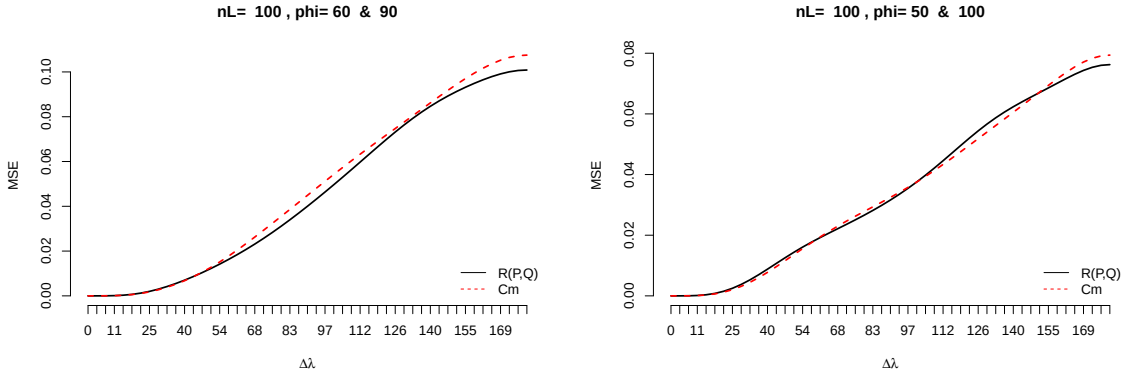


(a) Model 2 (4.7)



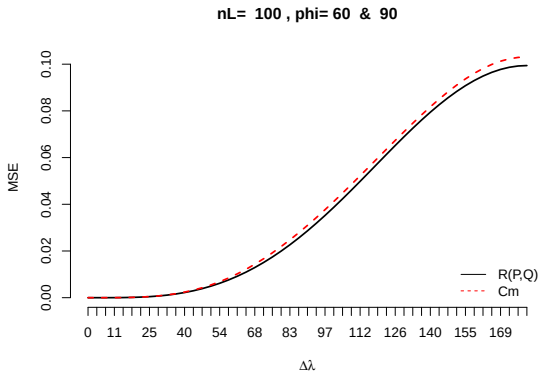
(b) Model 3 (4.8)

Figure 14. Cross covariance comparison of model 2 and model 3

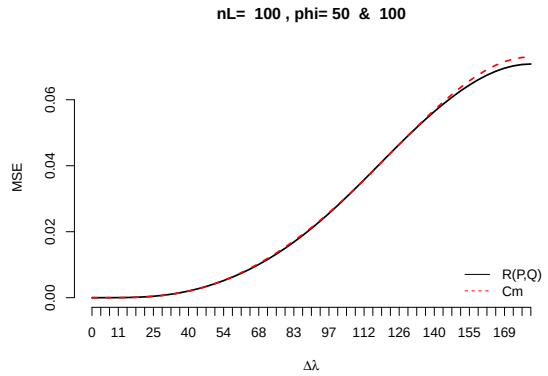


(a) Model 1 (pair 1)

(b) Model 1 (pair 2)



(c) Model 2 (pair 1)



(d) Model 2 (pair 2)

Figure 15. MSE comparison between C_m and $R(P,Q)$ using 500 simulations; pair 1 $(30^0S, 0^0)$, pair 2 $(40^0S, 10^0N)$ figures (a) - (b) is the comparison for model 1 and figure (c)-(d) is the comparison for model 2

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

		Set 1		Set 2	
Model	(ϕ_P, ϕ_Q)	$R(P,Q)$	C_m	$R(P,Q)$	C_m
Model1	(60, 90)	2.298 (0.0022)	2.427 (0.0004)	15.688 (0.0250)	16.384 (0.0015)
	(50, 100)	1.784 (0.0009)	1.782 (0.0001)	13.295 (0.0193)	12.767 (0.0030)
	(10, 150)	0.564 (0.0009)	0.623 (0.0001)	8.062 (0.0226)	9.177 (0.0023)
Model2	(60, 90)	2.000 (0.0004)	2.080 (0.0004)	12.452 (0.0042)	13.021 (0.0023)
	(50, 100)	1.437 (0.0001)	1.459 (0.0000)	9.196 (0.0015)	9.262 (0.0006)
	(10, 150)	0.457 (0.0014)	0.512 (0.0001)	6.034 (0.0337)	7.266 (0.0026)

One can see from the above table that a variety of pairs of latitudes and two different sets of parameters are used in simulations. The MSEs from C_m are comparable with those from direct $R(P,Q)$ approach while the C_m approach gives smaller bias.

5.4.5 Generated Data

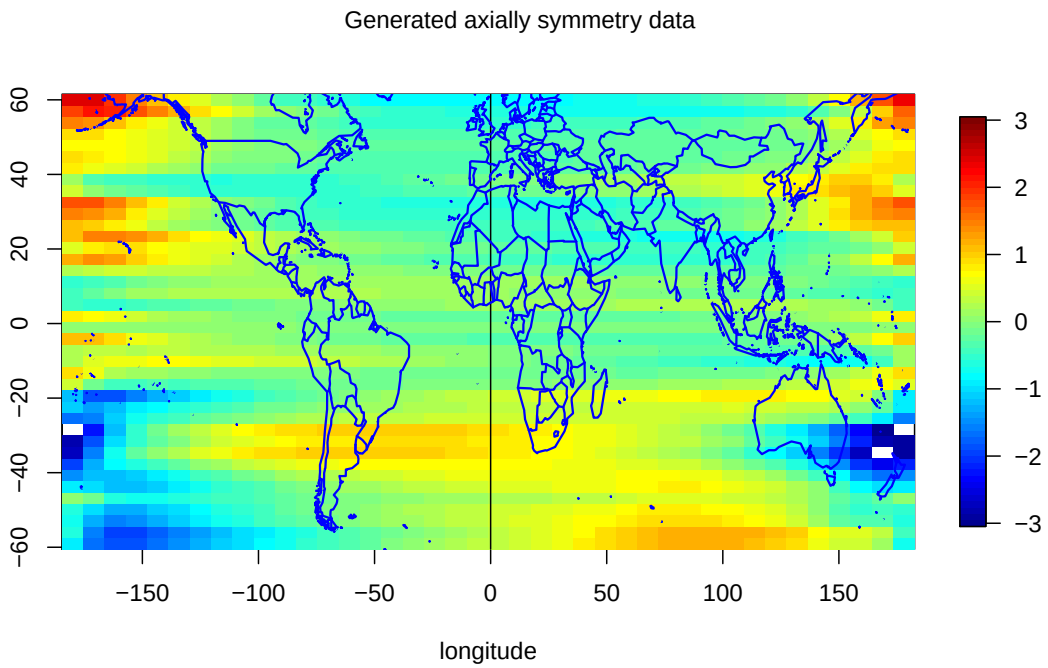


Figure 16. A snapshot of global data generated based on C_m approach using zero mean random process (model 2)

The Figure 16 is a snapshot of the global data generated based on model 2 and could potentially be used later for research. Clearly there are spatial trends within the latitudes but not within longitudes. It is somewhat difficult to use the covariance structure to generate data when it is closer to Earth's pole (similar complexity can also be observed in MSU and TOMS data). Therefore we produced a snapshot by generating 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).

However we observed some inconsistencies (strong spots) when examining closer close to the boundary points of longitudes ($\lambda \to \pm\pi$).

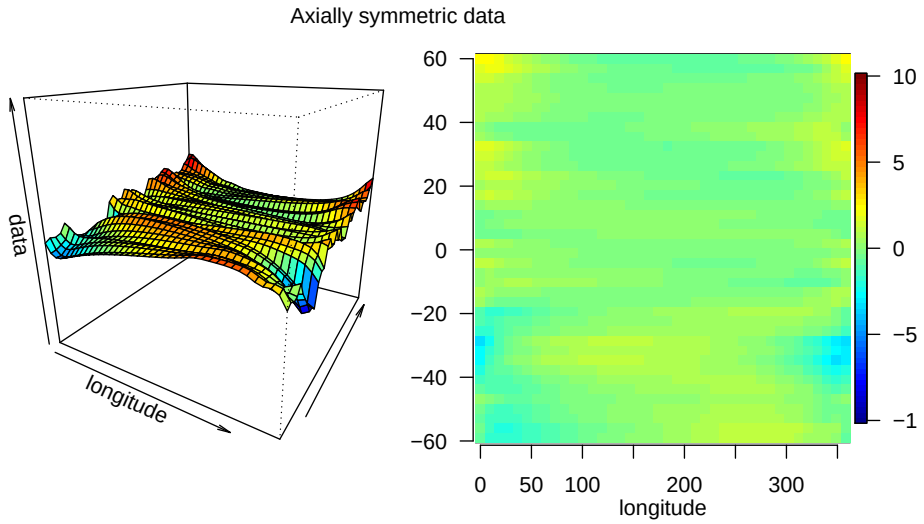


Figure 17. One snapshot of the axially symmetric data generated based on model 2, grid resolution $2^0 \times 1^0$ (data scale -10 and 10).

The above Figure 17 refers to the data snapshot given by Figure (16) and it is clear that trends are within latitudes not within longitudes. Four snapshots of the gridded data generated based on all models are given in Appendix A.

5.5 Discussion

The data generation algorithm proposed in this dissertation seems producing axially symmetric data that follow the given covariance models. The bias of the cross variogram estimates from the data generated based on our algorithm is generally smaller than that from direct $R(P, Q)$ approach, while maintaining the same level of MSEs. The computation cost for our algorithm is very small, with the order of

$O(n_L n_l^2)$. This is much smaller than that when taking the square root of $R(P,Q)$ through SVD decomposition. Note that the dimension of $R(P,Q)$ is $n_l n_L \times n_l n_L$, which might be expensive or even not possible when performing the SVD with large dimension. In addition, obtaining $R(P,Q)^{1/2}$ with the light of block circulant matrix seems unclear. Li (2013) indicated that one might find the eigenvalues using the properties of block circulant matrices, more specifically, they pointed out these eigenvalues are related to sub-matrices $R_0, R_1, \ldots, R_{n_L-1}$. However, since these matrices are not necessary symmetric, their eigenvalues could be complex-valued or negative. When submatrices are symmetric, Tee (2005) proposed some methods to find the eigenvalues. This will be extensively explored in the future.

CHAPTER VI

FUTURE RESEARCH

In this dissertation, we focus on data generation and estimation for axially symmetric processes on the sphere. We first show that for the stationary random process on the circle, the commonly used covariance function estimator based on MOM is biased with non-estimable bias, while the unbiased MOM variogram estimator is inconsistent. Our second project emphasizes on data generation, in which the axially symmetric random process 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 that generates axially symmetric data with the given covariance model. All of the above results and theories have been supplemented via simulations.

We can extend this dissertation work to a number of future research areas. We will first explore the unbiasedness and consistency of the MOM covariance and variogram estimators for homogeneous and axially symmetric random processes on the sphere. In particular, we expect a similar result as the one for circle holds for axially symmetric random processes. On the other hand, we will also explore the ergodic condition (if exists) that ensures the consistency of these estimators.

We have noticed that our proposed data generation algorithm assumes the closed form of $C_m(\phi_P, \phi_Q)$, which sometimes may not be available. This might restrict

the applicability of our data generation algorithm. Note that in order to implement the algorithm, we only need the $C_m(\phi_P, \phi_Q)$ over the gridded locations. Therefore, given the covariance structure $R(P, Q)$, we may use the Discrete Fourier Transform to obtain those gridded $C_m(\phi_P, \phi_Q)$ values. This would definitely complement our dissertation research.

Kriging, or making predictions at unobserved locations, has always been one of the important applications of data modeling and analysis in spatial statistics. With the complexity and dimensionality of global data, it is highly demanded that practically useful parametric models with interpretable parameters would be available for geography and environmental scientists. As the continuation of this dissertation research, we wish to enhance the kriging techniques and make use of proposed global data generation methods to make global predictions with less dimensionality.

REFERENCES

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

Christy, J. R., Spencer, R. W., and Braswell, W. D. (2000). Msu tropospheric temperatures: Dataset construction and radiosonde comparisons. *Journal of Atmospheric and Oceanic Technology*, 17(9):1153–1170.

Cressie, N. (1985). Fitting variogram models by weighted least squares. *Journal of the International Association for Mathematical Geology*, 17(5):563–586.

Cressie, N. (1993). *Statistics for spatial data*. John Wiley & Sons.

Cressie, N. and Johannesson, G. (2008). Fixed rank kriging for very large spatial data sets. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(1):209–226.

De Mazancourt, T. and Gerlic, D. (1983). The inverse of a block-circulant matrix. *IEEE transactions on antennas and propagation*, 31(5):808–810.

Dufour, J.-M. and Roy, R. (1976). On spectral estimation for a homogeneous random process on the circle. *Stochastic Processes and their Applications*, 4(2):107 – 120.

Fuller, W. A. (2009). *Introduction to statistical time series*, volume 428. John Wiley & Sons.

Gallager, R. G. (2008). Circularly-symmetric gaussian random vectors. *preprint*, pages 1–9.

Gneiting, T. (2013). Strictly and non-strictly positive definite functions on spheres. *Bernoulli*, 19(4):1327–1349.

Hitczenko, M. and Stein, M. L. (2012). Some theory for anisotropic processes on the sphere. *Statistical Methodology*, 9(1):211–227.

Huang, C., Zhang, H., and Robeson, S. M. (2011). On the validity of commonly used covariance and variogram functions on the sphere. *Mathematical Geosciences*, 43(6):721–733.

Huang, C., Zhang, H., and Robeson, S. M. (2012). A simplified representation of the covariance structure of axially symmetric processes on the sphere. *Statistics & Probability Letters*, 82(7):1346–1351.

Jeong, J. and Jun, M. (2015). A class of matérn-like covariance functions for smooth processes on a sphere. *Spatial Statistics*, 11:1–18.

Jones, R. H. (1963). Stochastic processes on a sphere. *The Annals of mathematical statistics*, 34(1):213–218.

Jun, M. and Stein, M. L. (2007). An approach to producing space: Time covariance functions on spheres. *Technometrics*, 49(4):468–479.

Jun, M. and Stein, M. L. (2008). Nonstationary covariance models for global data. *The Annals of Applied Statistics*, pages 1271–1289.

Leadbetter, H. C.-M. (1967). Stationary and related stochastic processes. *New York*.

Li, T.-H. and North, G. R. (1997). Aliasing effects and sampling theorems of spherical random fields when sampled on a finite grid. *Annals of the Institute of Statistical Mathematics*, 49(2):341–354.

Li, Y. (2013). Non-parametric and semi-parametric estimation of spatial covariance function.

Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498.

Matheron, G. (1973). The intrinsic random functions and their applications. *Advances in Applied Probability*, 5(3):439–468.

Muir, S. T. (1920). *The theory of determinants*.

Schoenberg, I. J. (1942). Positive definite functions on spheres. *Duke Math. J.*, 9(1):96–108.

Stein, M. L. (1999). *Interpolation of spatial data: some theory for kriging*. Springer Science & Business Media.

Stein, M. L. (2007). Spatial variation of total column ozone on a global scale. *The Annals of Applied Statistics*, pages 191–210.

Tee, G. J. (2005). Eigenvectors of block circulant and alternating circulant matrices.

Trapp, G. E. (1973). Inverses of circulant matrices and block circulant matrices. *Kyungpook Mathematical Journal*, 13(1):11–20.

Wackernagel, H. (2013). *Multivariate geostatistics: an introduction with applications*. Springer Science & Business Media.

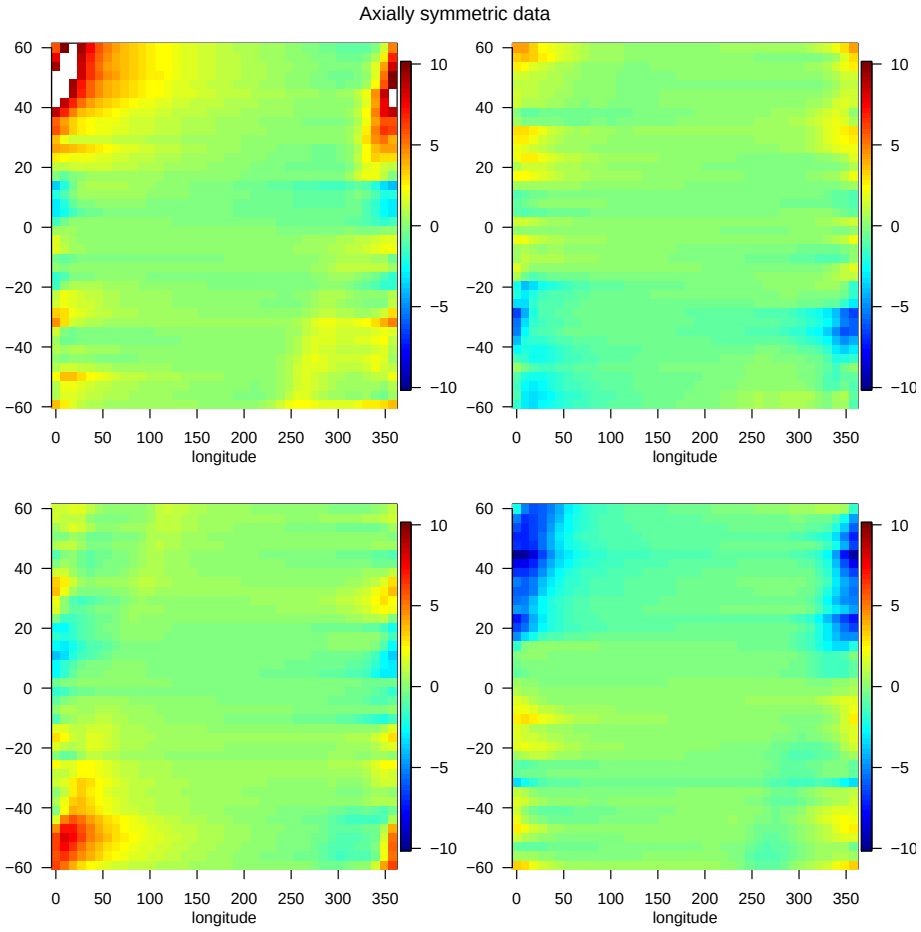
Yadrenko, M. I. (1983). *Spectral theory of random fields*. Optimization Software.

Yaglom, A. M. (1961). Second-order homogeneous random fields. In *Proc. 4th Berkeley Sympos. Math. Statist. and Prob*, volume 2, pages 593–622.

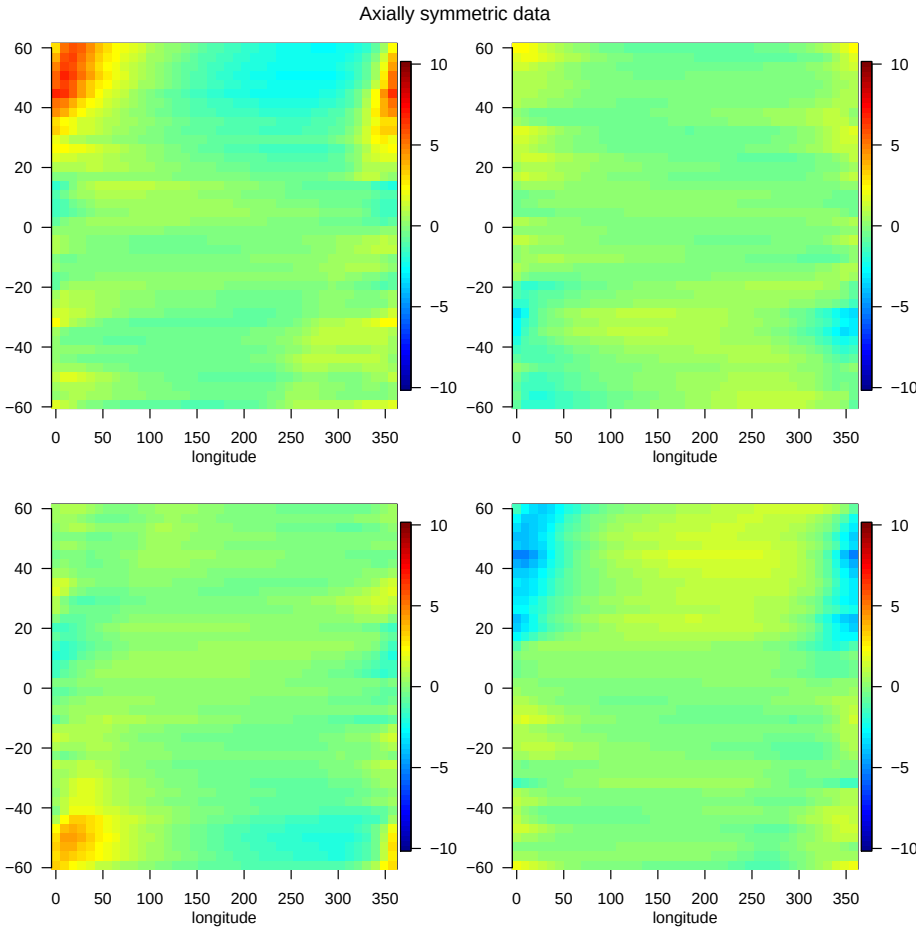
Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. 99(465):250–261.

APPENDIX A

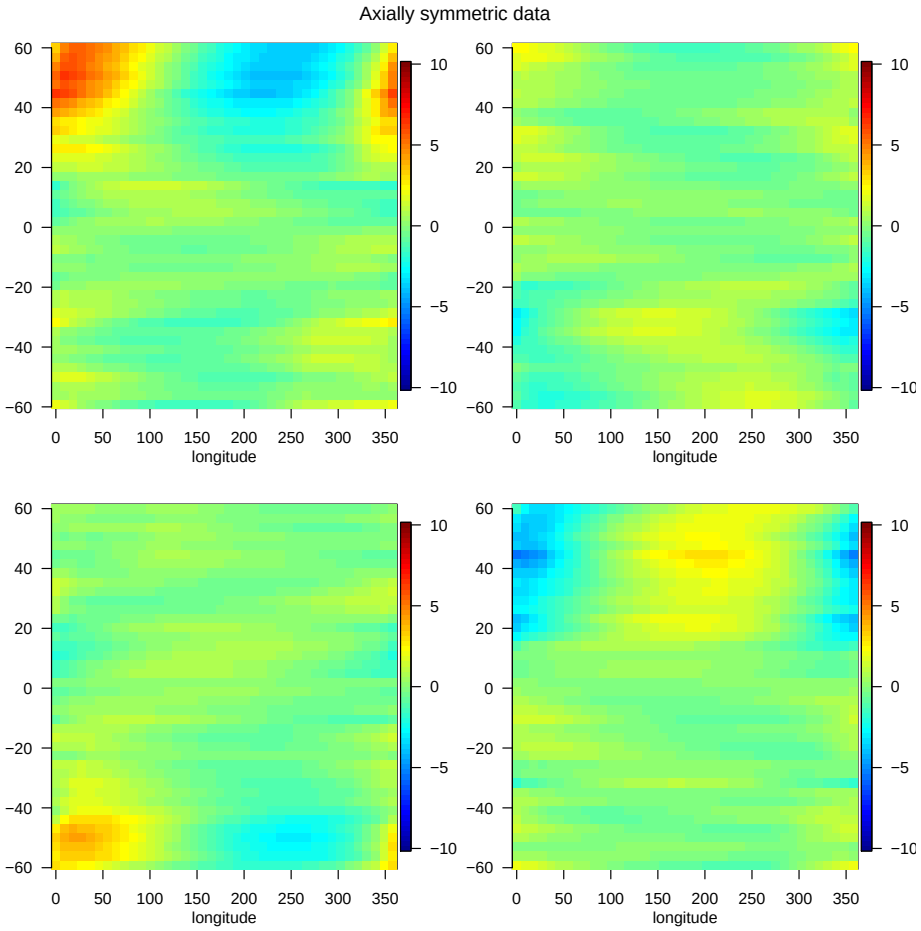
DATA SNAPSHOTS FOR ALL MODELS



Four consecutive axially symmetric data snapshots based on model 1 (4.6), grid resolution $2^0 \times 1^0$ (data scale -10 and 10).



Four consecutive axially symmetric data snapshots based on model 2 (4.7), grid resolution $2^0 \times 1^0$ (data scale -10 and 10).



Four consecutive axially symmetric data snapshots based on model 3 (4.8), grid resolution $2^0 \times 1^0$ (data scale -10 and 10).