Data Generation and Estimation for Axially Symmetric Processes on the Sphere

Chris Vanlangenberg[1]
Advisor: Dr. Haimeng Zhang[2]

University of North Carolina at Greensboro

September 13, 2016

[1]cdvanlan@uncg.edu
[2]h_zhang5@uncg.edu

Abstract

Review and update later

Global-scale processes and phenomena are of utmost importance in the geophysical sciences. Data from global networks and satellite sensors have been used to monitor a wide array of processes and variables, such as temperature, precipitation, etc. In this dissertation, we are planning to achieve explicitly the following objectives,

1. Develop both non-parametric and parametric approaches to model global data dependency.
2. Generate global data based on given covariance structure.
3. Develop kriging methods for global prediction.
4. Explore one or more of the popularly discussed global data sets in literature such as MSU (Microwave Sounding Units) data, the tropospheric temperature data from National Oceanic and Atmospheric and TOMS (Total Ozone Mapping Spectrometer) data, total column ozone from the Laboratory for Atmospheres at NASA's Goddard Space Flight Center Administration satellite-based Microwave Sounding Unit.

Global scale data have been widely studied in literature. A common assumption on describing global dependency is the second order stationarity. However, with the scale of the Earth, this assumption is in fact unrealistic. In recent years, researchers have focused on studying the so-called axially symmetric processes on the sphere, whose spatial dependency often exhibit homogeneity on each latitude, but not across the latitudes due to the geophysical nature of the Earth. In this research, we have obtained some results on the method of non-parametric estimation procedure, in particular, the method of moments, in the estimation of spatial dependency. Our initial result shows that the spatial dependency of axially symmetric processes exhibits both anti-symmetric and symmetric characteristics across latitudes. We will also discuss detailed methods on generating global data and finally we will outline our methodologies on kriging techniques to make global prediction.

Chapter 1

Introduction

In this chapter we have given a brief introduction to some of the basic concepts in spatial statistics which are necessary to follow other chapters in this dissertation. Moreover, we have discuss about stationarity, isotropy, intrinsic stationarity, covarince function and it properties, variogram, continuity and differentiability, spectral representations, Bochner's theorem, spectral densities, circulant matrices and it's properties with special cases.

1.1 Spatial random field

A real-valued spatial process in d dimensions or a spatial random field can be denoted as $\{Z(x) : x \in D \subset \mathbb{R}^d\}$ where x is the location of process $Z(x)$ and x varies over the set D which is fixed and discrete. The distribution of the random vector $Z(\underset{\sim}{X}) = (Z(x_1), \ldots, Z(x_n))$ is given by the associated finite-dimensional joint distributions

$$F\{Z(x_1), \ldots, Z(x_n)\} = P\{Z(x_1) \leq h_1, \ldots, Z(x_n) \leq h_n\} \tag{1.1.1}$$

A random process is a collection of random variables $X \in \{X(s) : s \in D\}$, defined in a common probability space. In general, if

- $s \in N$: $X(s)$ is a random sequence which is used in time series.
- $s \in R^1$: $X(s)$ is a random process which is also referred as a stochastic process.
- $s \in R^d$: $X(s)$ is a random filed or a spatial process if $d > 1$
- $s \in S^2$: $X(s)$ is a random process on the sphere.
- $s \in R^d \times R$: $X(s)$ is a spatio-temporal process which involves location and time.

? comment : more words

1.1.1 Stationarity and Isotropy

A spatial random field is strict stationarity, for all finite n, $x_1, \ldots, x_n \in \mathbb{R}^d$, $h_1, \ldots, h_n \in \mathbb{R}$ and $x \in \mathbb{R}^d$, if the random field is invariant under translation. that is,

$$P\{Z(x_1 + x) \leq h_1, \ldots, Z(x_n + x) \leq h_n\} = P\{Z(x_1) \leq h_1, \ldots, Z(x_n) \leq h_n\} \tag{1.1.2}$$

Strict stationarity is a too strong condition as it involves the distribution of the random field but many spatial methods are based moments. Therefore, it is sufficient to use weak assumptions and we could say a random process $Z(x)$ is weakly stationary if,

$$\begin{aligned} E(Z(x)) &= \mu \\ E^2(Z(x) &< \infty \\ C(h) &= Cov(Z(x), Z(x+h)) \end{aligned} \tag{1.1.3}$$

if $Z(x)$ has a finite second moment with constant mean and $C(h)$ the covarince (also referred as auto-covariance) function depends on the spatial distance only. Further strictly stationary random fields with finite second moment is also weakly stationary, but weak stationarity does not imply strict stationarity. However, in the case of Gaussian random fields that weakly stationary are also strict stationarity because the first two moments (μ, σ) will explain the distribution.

Suppose $Z(x)$ is weakly stationary on $\mathbb{R}^d$ with autocavariance function $C(h)$ then it has the following properties,

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

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

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

(iv) If C_1, C_2, C_n are valid covariance functions then following $C(x)$ functions are also valid covarince functions

  (a) $C(x) = a_1 C_1 + a_2 C_2$, $\forall a_1, a_2 \geq 0$
  (b) $C(x) = C_1(x) C_2(x)$
  (c) $\lim\limits_{n \to \infty} C_n(x) = C(x)$, $\forall x \in \mathbb{R}^d$

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

$$\sum_{i,j=1}^{N} a_i a_j C(x_i - x_j) \geq 0, \tag{1.1.4}$$

for any integer N, any constants $a_1, a_2, \ldots, a_N$, and any locations $x_1, x_2, \ldots, x_N \in \mathbb{R}^d$. Positive definiteness is a necessary and sufficient condition to have a valid covariance function.

Theorem 1.1.1 (Mercer's theorem (simplified version))

A kernal $K : [a, b] \times [a, b] \to \mathbb{R}$ be a symmetric continuous function that is non-negative definite,

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j K(s,t) \geq 0 \quad \textit{and} \quad K(s,t) = K(t,s)$$

for all $(s,t) \in [a,b]$ and $a_i > 0$. Let $T_K : L_2 \to L_2$ be an intergral operator defined by

$$[T_K f](\cdot) = \int_a^b K(\cdot, s) f(s) ds$$

is positive, for all $f \in L_2$

$$\int_a^b K(s,t) f(s) f(t) ds dt \geq 0.$$

The corresponding orthonormal eigen functions $\psi_i \in L_2$ and non negative eigen values $\lambda_i \geq 0$ of the operator T_K is defined as

$$T_K(\psi_i(\cdot)) = \int K(\cdot, s)\psi(s) ds = \lambda_i \psi_i(\cdot), \quad \int \psi_i(\cdot)\psi_j(\cdot) = \delta_{ij}$$

then the kernal $K(\cdot)$ is a uniformly convergent series in terms of eigen functions and associated eigen values of T_K as follows,

$$K(s,t) = \sum_{j=1}^{\infty} \lambda_i \psi_i(s)\psi_i(t)$$

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

$$P\{Z(Hx_1 + x) \leq h_1, \ldots, Z(Hx_n + x) \leq h_n\} = P\{Z(x_1) \leq h_1, \ldots, Z(x_n) \leq h_n\} \quad (1.1.5)$$

Isotropy assumes that it is not required to distinguish one direction from another for the random field $Z(x)$.

If the variance between two locations solely depends on the distance between the two locations then the process is said to be intrinsically stationary. Semivariogram is an alternative to the covariance function proposed by Matheron. For an intrinsically stationary random field $Z(x)$,

$$\begin{aligned} E[Z(s)] &= \mu, \\ \gamma(h) &= \frac{1}{2} Var(Z(s+h) - Z(x)), \end{aligned} \quad (1.1.6)$$

Where γ is the semivariogram and $\gamma(h) = C(0) - C(h)$ for a weakly stationary process with covariance function $C(h)$. Intrinsic stationary is defined in terms of variogram and it is more general than weak stationary which is defined in terms of covariance. Clearly, when $C(h)$ is known we can get $\gamma(h)$ but not $C(h)$ when $\gamma(h)$ is known. For example consider a linear semi variogram function,

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

when $\lim_{h\to\infty} \gamma(h) \to \infty$ thus this is not weak stationary and $C(h)$ does not exist.

Similar to covariance the variogram is conditionally negative definite if only if

$$\sum_{i,j=1}^{N} a_i a_j 2\gamma(x_i - x_j) \leq 0, \tag{1.1.7}$$

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

spectral representation of variogram

1.1.2 Mean square continuity & differentiability

There is no simple relationship between $C(h)$ and the smoothness of $Z(x)$. 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 can say, $\{X_n\}$ converges in L^2 if there exists such a 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 weak stationary and $C(\cdot)$ is the covariance function then $E(Z(x+h) - Z(x))^2 = 2(C(0) - C(h))$. Therefore $Z(x)$ is mean square continuous iff $C(\cdot)$ is continuous at the origin.

1.1.3 Circularly-symmetry Gaussian random vectors

Sometimes it is convenient to use complex valued random functions, rather than real valued random functions.

We say, $Z(x) = U(x) + iV(x)$ is a complex random field if $U(x), V(x)$ are real random fields. If $U(x), V(x)$ are weakly stationary so does $Z(x)$.The covariance function can be defined as,

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

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

$$\sum_{i,j=1}^{n} c_i \bar{c_j} C(x_i - x_j) \geq 0 \tag{1.1.8}$$

In general, a normal family has two parameters, location parameter μ and scale parameter Σ. But when we are dealing with complex normal family there is one additional parameter, the relation matrix also referred as pseudo-covariance matrix. In the case of real normal family the pseudo-covariance matrix is equivalent to covariance matrix.

According to Gallager (2008), a complex random variable $Z = Z^{Re} + iZ^{Im}$ is Gaussian, if Z^{Re}, Z^{Im} both are real and jointly Gaussian. Then Z is circularly-symmetric if both Z and $e^{i\phi}Z$ has the same probability distribution for all real ϕ. Since $E[e^{i\phi}Z] = e^{i\phi}E[Z]$, any circularly-symmetric complex random vector must have $E[Z] = 0$, in other words its mean must be zero.

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

$$K_{\underset{\sim}{Z}} = E[\underset{\sim}{Z}\underset{\sim}{Z}^*] \tag{1.1.9}$$

$$M_{\underset{\sim}{Z}} = E[\underset{\sim}{Z}\underset{\sim}{Z}^T] \tag{1.1.10}$$

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

For example, consider a vector $\underset{\sim}{Z} = (Z_1, Z_2)^T$ where $Z_1 = Z_1^{Re} + iZ_1^{Im}$ and $Z_2 = Z_1^*$ ($Z_2^{Re} = Z_1^{Re}, Z_2^{Im} = -Z_1^{Im}$). The four real and imaginary parts of $\underset{\sim}{Z}$ are jointly Gaussian (each $N(0, 1/2)$), so $\underset{\sim}{Z}$ is complex Gaussian.

Now, the covariance and pseudo-covariance matrices are different defined by 1.1.9 and 1.1.10 respectively given by

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

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

It is easy to notice that $E[Z_1^2] = E[z_1^{Re} z_1^{Re} - z_1^{Im} z_1^{Im}] = 1/2 - 1/2 = 0$ and if Z_1 is real (obviously also Z_2) then covarince and pseudo-covariance matrices are the same *i.e.* $M_Z \equiv K_Z$

The covariance matrix of real $2n$ random vector $\underset{\sim}{Z} = (\underset{\sim}{Z}^{Re}, \underset{\sim}{Z}^{Im})^T$ is determined by both $K_{\underset{\sim}{Z}}$ and $M_{\underset{\sim}{Z}}$ as follows,

$$\begin{aligned} E[\underset{\sim}{Z}^{Re}\underset{\sim}{Z}^{Re}] &= \frac{1}{2}Re(K_{\underset{\sim}{Z}} + M_{\underset{\sim}{Z}}), \\ E[\underset{\sim}{Z}^{Im}\underset{\sim}{Z}^{Im}] &= \frac{1}{2}Re(K_{\underset{\sim}{Z}} - M_{\underset{\sim}{Z}}), \\ E[\underset{\sim}{Z}^{Re}\underset{\sim}{Z}^{Im}] &= \frac{1}{2}Im(-K_{\underset{\sim}{Z}} + M_{\underset{\sim}{Z}}), \\ E[\underset{\sim}{Z}^{Im}\underset{\sim}{Z}^{Re}] &= \frac{1}{2}Im(K_{\underset{\sim}{Z}} + M_{\underset{\sim}{Z}}) \end{aligned} \tag{1.1.11}$$

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

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

Theorem 1.1.2 (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.

1.1.4 Spectral representation of a random field

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

$$Z(x) = \sum_{k=1}^{n} Z_k e^{i\omega_k^T x}. \tag{1.1.12}$$

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

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

$$C(x) = \int_{\mathbb{R}^d} e^{i\omega^T x} F(d\omega) \tag{1.1.13}$$

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

Theorem 1.1.3 (Bochner's Theorem)

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

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

$$f(\omega) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\omega^T x} C(x) dx \tag{1.1.14}$$

1.1.5 Spectral densities

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

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

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

$$C(h) = \frac{\pi^{1/2}\phi}{2^{\nu-1}\Gamma(\nu + 1/2)\alpha^{2\nu}} (\alpha|h|)^\nu Y_\nu(\alpha|t|) \tag{1.1.15}$$

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

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

1.2 Circulant matrix

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

$$A = circ[a_o, a_1, \cdots, a_{n-1}] \quad = \quad \begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{bmatrix}. \tag{1.2.1}$$

The eigenvalues of A are given by

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

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

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

If matrix A is real symmetric then its eigen values are real; for even $n = 2h$ the eigen values $\lambda_j = \lambda_{n-j}$ or there are either two eigen values or none with odd multiplicity, for odd $n = 2h - 1$ the eigen value λ_0 equal to any λ_j for $1 \leq j \leq h - 1$ 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 eigen vectors with corresponding real eigen values.

1.2.1 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,

$$B = bcirc[a_o, a_1, \cdots, a_{n-1}] = \begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{bmatrix}. \tag{1.2.2}$$

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 $(np \times np)$ identity matrix and the eigen values of M given by ρ_l, the eigen matrix of M can be given by $Q_{np\times np} = \{\underset{\sim}{\psi_0}, \underset{\sim}{\psi_1}, \ldots, \underset{\sim}{\psi_{n-1}}\}$. From Trapp (1973) it can be shown that $Q^{-1} = Q^*/n$ where Q^* is the conjugate transpose of Q now we can write,

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

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

$$d_i = \sum_{k=0}^{n-1} a_k e^{i2lk\pi/n}$$

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

$$B^{-1} = Q \cdot \begin{pmatrix} d_0^{-1} & 0 & \cdots & 0 \\ 0 & d_1^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n{-}1 \end{pmatrix} \cdot Q^{-1}.$$

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

When $a_{j_{1\times 1}}$, $B = A$, $d_i^{-1} = \lambda^{-1}$, and the eigen matrix 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 real also symmetric and $Q^{-1} = Q^T$.

should we add the following special cases?

Case 1 : When a_j's are symmetric (*by Tee, add citation*)
Case 2 : When a_j's are circulant

Chapter 2

Literature Review (due August 28)

Chapter 3

Covariance and Variogram Estimation on the Cricle

3.1 Stationary process on a circle

In this chapter we consider a real valued process $\{X(P) : P \in S\}$ on the unit circle S, with finite second moment and continuity in quadratic mean. According to Dufour and Roy (1976) the process $\{X(P)\}$ can be represented in a Fourier series which is convergent in quadratic mean,

$$X(P) = A_0 + \sum_{n=1}^{\infty}(A_n \cos(nP) + B_n \sin(nP). \tag{3.1.1}$$

$$\begin{aligned} \text{where} \quad A_0 &= \frac{1}{2\pi}\int_S X(P)dP, \\ A_n &= \frac{1}{\pi}\int_S \cos(nP)dP \\ B_n &= \frac{1}{\pi}\int_S \sin(nP)dP \end{aligned} \tag{3.1.2}$$

Let P, Q be any two points on the circle, the covariance $R(P, Q)$ between the two points can be defined as,

$$R(P, Q) = E(X(P)X(Q)) = cov(X(P), X(Q))$$

The process $X(P)$ is stationary if $E(X(P))$ is a constant and $R(P, Q)$ is function of the angular distance θ_{PQ} between P and Q. If the process $X(P)$ is stationary on the circle,

$$cov(A_n, A_m) = a_n\delta(n, m) = cov(B_n, B_m), \quad \text{for } n, m \geq 0. \tag{3.1.3}$$

Let $\{X(t_k)\}$ be a collection of gridded observations on a circle, with $t_k = (k-1)*2\pi/n, k = 1, 2, \cdots, n$. Let $C(\theta), \theta \in [0, \pi]$ denote a stationary covariance function on the circle, the

underlying process is stationary if it's covariance function solely depends on the distance θ and given as follows,

$$C(\theta) = cov(X(t_k + \theta), X(t_k)), \qquad \theta \in [0, \pi]. \tag{3.1.4}$$

The above covariance function can we can be written as a Fourier summmation (Roy (1972))

$$C(\theta) = \sum_{n=0}^{\infty} a_n \cos(n\theta), \tag{3.1.5}$$

with $\sum_{n=0}^{\infty} a_n < \infty$, and $a_n \geq 0$. Note that

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

3.2 Estimation

Assuming the covariance function $C(\theta)$ on a cricle is a continuous function on $[0, \pi]$) and the gridded points $\{X(t_k)\}$ on a circle can be represented as $\underset{\sim}{X} = (X_1, X_2, \cdots, X_n)^T$. The variance-covariance matrix of the sample vector $\underset{\sim}{X}$ is given by Σ. Lets assume $E(X(t)) = \mu$ is unknown, and estimated by $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X(t_i)$. Further, we can denote $\bar{X}$ in the following form,

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

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

$$\begin{aligned} \text{When } n \to \infty, \frac{1}{n} \left(2 \sum_{m=0}^{N} C(m2\pi/n) \right) &= \frac{1}{\pi} \frac{\pi}{N} \left(\sum_{m=0}^{N} C(m2\pi/n) \right) \\ &\to \frac{1}{\pi} \int_0^{\pi} C(\theta) d\theta = a_0. \end{aligned}$$

$$\begin{aligned} \text{Hence} \quad var(\bar{X}) &= \frac{1}{n} \left(2 \sum_{m=0}^{N} C(m2\pi/n) \right) - \frac{1}{n}(C(0) + C(\pi)) \\ &\to a_0, \qquad \text{as } n \to \infty. \end{aligned}$$

We can conclude that $var(\bar{X}) \to \frac{1}{\pi}\int_0^{\pi} C(\theta)d\theta \quad \text{as } n \to \infty.$

Proposition 3.2.1 *Let $\bar{X}$ be an unbiased estimator for μ then $\bar{X}$ is not a consistent estimator for μ, mean on a circle.*

proof: *If $\bar{X}$ is consistant we get the following,*

$$P(|\bar{X} - \mu| > \varepsilon) \to 0.$$

If $var(\bar{X}) \to 0$ and from Chebyshev's inequality we have

$$P(|\bar{X} - \mu| > \varepsilon) \leq \frac{var(\bar{X})}{\varepsilon^2} \to 0, \quad \text{for any } \varepsilon > 0.$$

Therefore, $var(\bar{X}) \to 0$ is a sufficient condition for consistency, but it is not necessary. However, if we assume $\underset{\sim}{X}$ is multivariate normally distributed, then $\bar{X}$ follows normal distribution with mean μ and approximate variance σ_0. Then ($Z \sim N(0,1)$)

$$\begin{aligned} P(|\bar{X} - \mu| > \varepsilon) &= P\left(\frac{|\bar{X} - \mu|}{\sqrt{a_0}} > \frac{\varepsilon}{\sqrt{a_0}}\right) \\ &= P\left(|Z| > \frac{\varepsilon}{\sqrt{\sigma_0}}\right) \not\to 0 \end{aligned}$$

since $\frac{\varepsilon}{\sqrt{\sigma_0}}$ is a fixed constant for each fixed $\varepsilon > 0$.

3.2.1 Estimation of covariance on a circle

We used method of moments (MOM) to estimate the covariance $C(\theta)$ on a circle, the estimator can be given by

$$\hat{C}(\Delta\lambda) = \frac{1}{n}\sum_{i=1}^{n}(X(t_i + \Delta\lambda) - \bar{X})(X(t_i) - \bar{X}), \tag{3.2.1}$$

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

We can show that the above estimator is a biased estimator for $C(\theta)$.

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

Remark 1 The MOM estimator $\hat{C}(\Delta\lambda)$ of the covariance function is actually a biased estimator with the shift amount of a_0. Therefore, if $a_0 = 0$ for a covariance function, we have the unbiased estimator $\hat{C}(\Delta\lambda)$.

Remark 2 If the gridded points were on a line, for example in time series, $E(\bar{X}-\mu)^2 \to 0$ as $n \to \infty$ under the assumption that the covariance function $C(\theta) \to 0$ when $\theta \to \infty$ (which is practically feasible), that is, $\bar{X}$ is consistent in the case of points on a line. In the case of circle, we might not have $C(\theta)$ close to 0 since θ is within a bounded region ($(0,\pi)$ for the circle) and we normally assume $C(\theta)$ is continuous for θ.

3.2.2 Estimation of variogram on a circle

The theoretical variogram function is given by,

$$\gamma(\theta) = C(0) - C(\theta). \tag{3.2.2}$$

and the MOM estimator for the variogram is given by,

$$\hat{\gamma}(\Delta\lambda) = \frac{1}{2n}\sum_{i=1}^{n}(X(t_i+\Delta\lambda)-X(t_i))^2. \tag{3.2.3}$$

We can show that variogram estimator through MOM is an unbiased estimator,

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

need to prove consistency

3.3 Data generation on a circle

First, we will discuss how to generate correlated data at n girdded points on a circle when the covarince function is defined and compare above covariance and variogram estimators. Since the observed data are correlated, the covariance function can be written as a function of distance (angle). For the data generation process we will use exponential family and power family covariance function as given below,

$$C(\theta) = C_1 e^{-a|\theta|} \quad a>0, C_1>0 \tag{3.3.1}$$

$$C(\theta) = c_0 - (|\theta|/a)^{\alpha} \quad a>0, \alpha\in(0,2] \text{ and } c_0 \geq \int_0^{\pi} (\theta/a)^{\alpha}\sin\theta d\theta \tag{3.3.2}$$

where $\theta = i*\Delta\lambda = \pm i*2\pi/n, i=1,2,\cdots,\lfloor n/2\rfloor$

Clearly, each location is correlated with other $n-1$ locations and $C(\theta)=C(-\theta)$ the variance-covariance matrix Σ is circulant and will be in the following form,

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

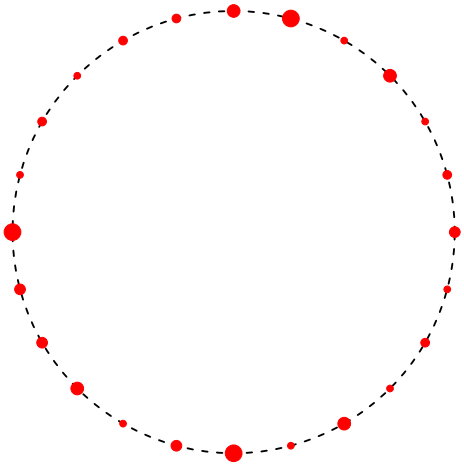

Figure 3.1: Random process on a circle at 24 points ($\Delta\lambda = 15^0$), the red dots represent the observed values at a given time and each point is associated with a random process of it's own.

where $\delta = 2\pi/n$, $\Lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ and $Q = \{\psi_1, \psi_2, \cdots, \psi_n\}$ are the respective eigen values and eigen vectors of the above circulant matrix. Now using singular value decomposition (SVD) we can obtain the correlated data $\underset{\sim}{X}$ on a circle as follows,

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

where $Z \sim N(\underset{\sim}{0}, 1_n)$.

3.3.1 Compare covariance estimator

In section 3.2.1 we proved that, in general the covarince estimator (3.2.1) on a circle is biased, with a bias of $var(\bar{X})$. In order to make things simple we set $C_1, a = 1$ when $\alpha = 0.5$ $c_0 \geq \int_0^\pi (\theta)^{0.5} \sin\theta d\theta$, from Fresnel intergal it can be shown that $c_0 \geq 2.4353$. Now we compare the covariance estimator to it's theoretical covarince given by equations 3.3.1 and 3.3.2. We computed the MOM estimator $\hat{C}(\theta)$ with 48 gridded observations on the circle from 500 simulations.

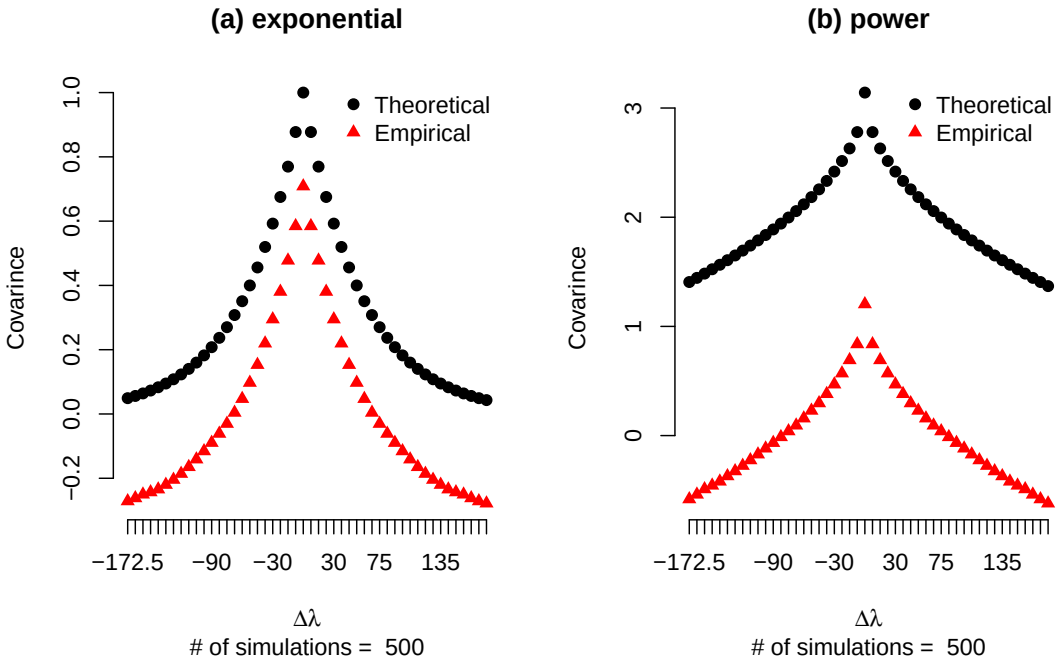

Figure 3.2: Theoretical and empirical covariance (with bias) comparison on a circle

Remark 1: We noticed that the shift between theoretical and empirical values were approximately equal to a_0 and we can obtain a_0 for the above covarince as follows,

$$\text{exponential}: \; a_0 = \frac{C_1}{a\pi}(1 - e^{-a\pi})$$

$$\text{power}: \; a_0 = c_0 - \left(\frac{\pi}{a}\right)^{\alpha} \frac{1}{\alpha + 1}$$

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

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

If the new covariance function $D(\theta)$ was used to generate the data on a circle then the theoretical and empirical values will match perfectly.

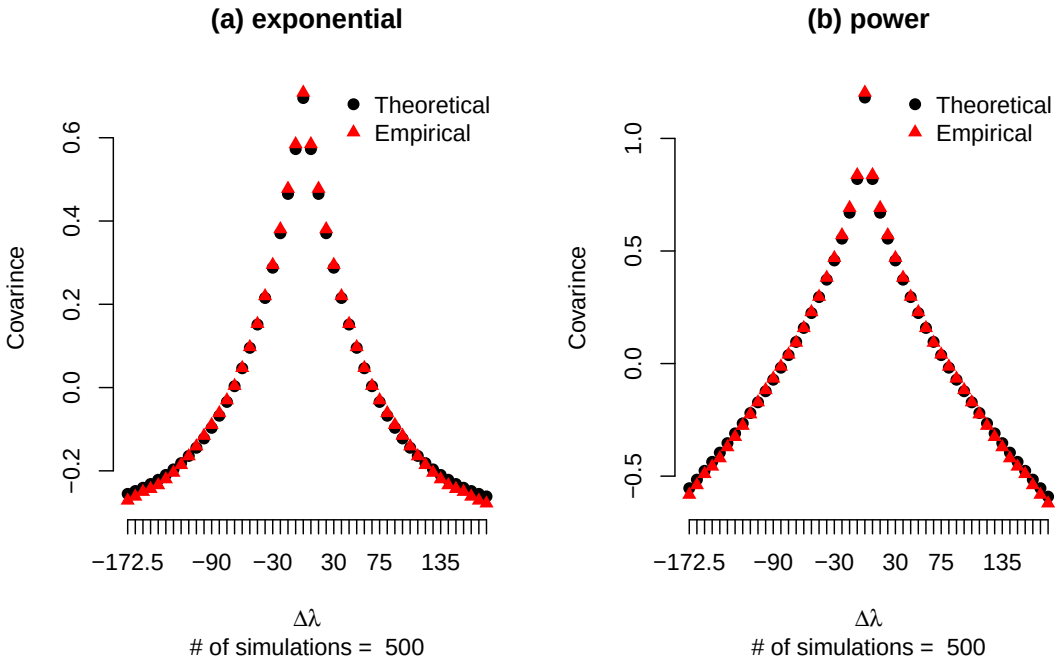

Figure 3.3: Theoretical and empirical covariance comparison on a circle

Remark 2: If the process is a zero mean process the covariance estimator will be unbiased ($i.e. Var(\bar{X}) = 0$) hence we will get a perfect match between theoretical and empirical values.

Remark 3: We have shown that covariance estiamtor is biased and the biasness will approach to a_0. When covariance function is unknown the biaseness a_0 is also known and the biasness cannnot to estimated (cannot find the variance of $\bar{X}$) from one circle, however in both exponential and power covariance models it can be estimated if multiple copies of data (on a circle) were generated *i.e.* $\hat{a} = var(\bar{X})$.

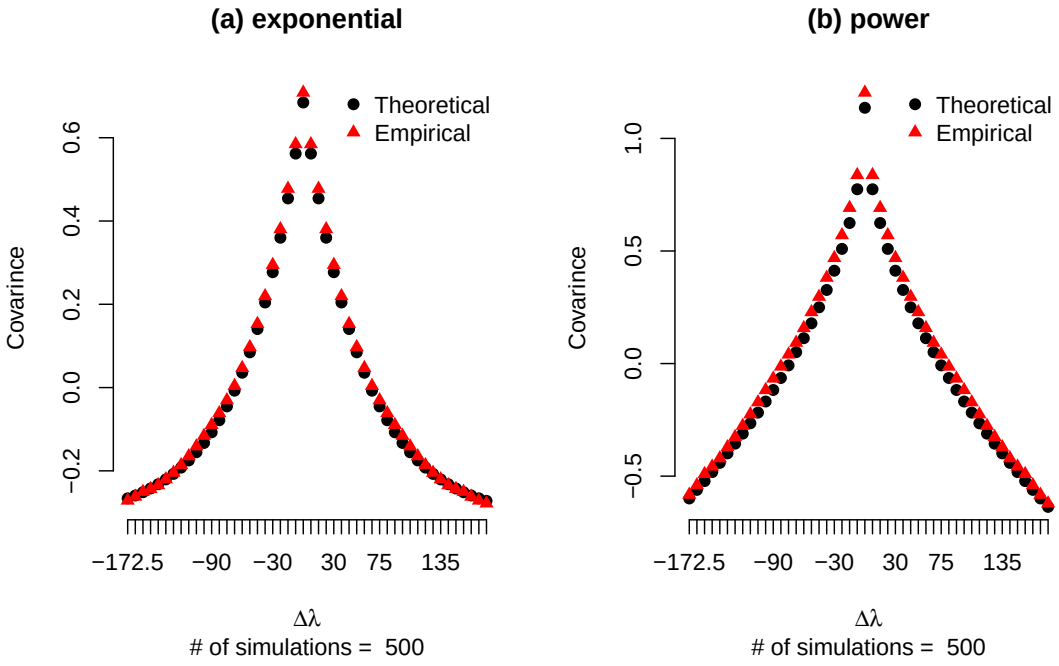

Figure 3.4: Theoretical and empirical covariance comparison on a circle, after removing biasness

3.3.2 Compare variogram estimator

We proved that in general the variogram estimator is unbiased *and consistent*. Since the semi variogram $\gamma(\theta) = C(0) - C(\theta)$, the theoretical variogram based on exponential and power covariance functions can be given in the following form,

$$\text{exponential}: \ \gamma(\theta) = C(0) - C(\theta) = C_1(1 - e^{-a|\theta|})$$

$$\text{power}: \ \gamma(\theta) = C(0) - C(\theta) = C_1(1 - e^{-a|\theta|})$$

We computed the variogram estimator $\hat{\gamma}(\theta)$ with 48 gridded observations on the circle from 500 simulations and there is a better fit between theoretical and empirical values compared to covariance models.

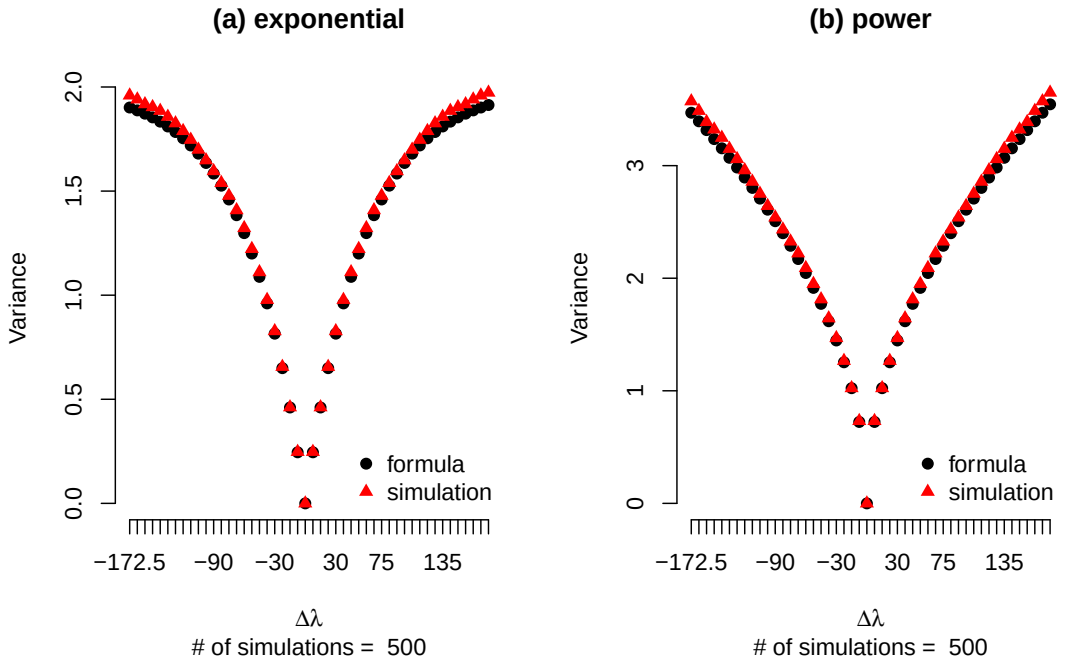

Figure 3.5: Theoretical and empirical comparison for variogram on a circle

Remark 4 *should we talk about how n_L is related with number of simulations*

Chapter 4

Random Process on a Sphere

Suppose $X \in \{X(s) : s \in D\}$, defined in a common probability space $s \in S^2$: $X(s)$ is a random process or a spatial processes on the sphere S^2 with radius r and s represents any location on sphere by latitude L and longitude l, where $0 \geq L \geq \pi$ and $-\pi \geq l \geq \pi$.

4.1 Covariance on sphere

The random process $X(\cdot)$ is said to be homogeneous (or isotropy) on S^2 if second moment is finite and invariant under the rotations on the sphere with constant mean. Similarly, we can define an isoropic random process on a sphere as,

$$\begin{aligned} E(X(s)) &= \mu \quad \text{for any } s \in S^2 \\ Cov(X(s_1), X(s_2)) &= C(\theta_{s_1 s_2}) \end{aligned}$$

where $\theta_{s_1 s_2}$ is the spherical angle between two locations s_1, s_2. For a unit sphere, the distance between the two locations can be defined as great circle distance ($\text{gcd}_{s_1 s_2}$) or chordal distance ($ch_{s_1 s_2}$) as follows,

$$\begin{aligned} \theta_{s_1 s_2} &= \arccos\left(\sin(L_1)\sin(L_2) + \cos(L_1)\cos(L_2)\cos(l_1 - l_2)\right) \\ gcd_{s_1 s_2} &= 1 \cdot \theta_{s_1 s_2} \\ ch_{s_1 s_2} &= 2\sin(\theta_{s_1 s_2}/2) \end{aligned}$$

In the case of $\mathbb{R}^d$, non-negative definite is a necessary and sufficient condition for a valid covariance function defined on $\mathbb{R}^d$ (1.1.4). Similarly, a real continuous function $C(\cdot)$ defined on the sphere is a valid covariance function if and only if it is non-negative definite,

$$\sum_{i,j=1}^{N} a_i a_j C(\theta_{s_1 s_2}) \geq 0, \tag{4.1.1}$$

for any integer N, any constants $a_1, a_2, \ldots, a_N$, and any locations $s_1, s_2, \ldots, s_N \in S^2$.

Let $P_k^{\nu}(\cos\theta)$ be the ultraspherical polynomials defined by the following infinite summation,

$$\frac{1}{(1-2c\cos\theta+c^2)^{\nu}} = \sum_{k=0}^{\infty} c^k P_k^{\nu}(\cos\theta) \quad \nu > 0 \tag{4.1.2}$$

$$\text{When } \nu = 0,\ P_k^0(\cos\theta) = \cos(k\theta)$$

Accoriding to Schoenberg (1942), a real continous function $C(\theta)$ is a valid covariance function on S^d, where $d = 1, 2, \ldots$, if and only if it can be written in the following form

$$C(\theta) = \sum_{k=0}^{\infty} c_k P_k^{(\nu)}(\cos\theta), \quad \nu = \frac{1}{2}(d-1) \tag{4.1.3}$$

where $\forall c_k \geq 0$ and $\sum c_k < \infty$.

Suppose $C(\cdot)$ is a covariance functions that is valid in S^d then it is valid on S^m where $d \leq m$. In general we have the following property,

$$S^1 \subset S^2 \subset \cdots S^d \subset \cdots S^{\infty}$$

$$C(S^1) \supset C(S^2) \supset \cdots C(S^d) \supset \cdots C(S^{\infty})$$

and covariance functions that are valid on S^m may not be valid on S^d where $m > d$ *example?*.

In chapter 3 we discussed about random processes on a circle and the covariance on a circle ($d = 1$) was expressed as follows,

$$C(\theta) = \sum_{k=0}^{\infty} c_k \cos(k\theta) \tag{4.1.4}$$

Since $\cos\theta \in S^1$ clearly $\cos\theta \in S^2$ and from the properties of covariance discussed in chapter 1, $P_k \cos\theta \in S^2$ where P_k is a Legender ploynomial. Therefore, when $d = 2$ the covarince on a sphere (S^2) can be expressed as follows,

$$C(\theta) = \sum_{k=0}^{\infty} c_k P_k(\cos\theta) \tag{4.1.5}$$

Since the Legendre polynomials are orthogonal we have

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

and on a sphere the coefficients c_k can be obtained by

$$c_{\nu} = \frac{2\nu+1}{2} \int_0^{\pi} C(\theta) P_{\nu}(\cos\theta) d\theta. \quad \nu = 0, 1, 2, \ldots \tag{4.1.6}$$

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

All covariance models that are valid on $\mathbb{R}^d$ are not valid on the sphere (S^2), Huang et al. (2011) evaluated the validity of commonly used covariance on a sphere that are valid on $\mathbb{R}^d$, they showed that some models are not valid on the sphere and some models are valid only for certain parameter values.

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

Table 4.1: Validity of covariance functions on the sphere, $a > 0, \theta \in [0, \pi]$. *When $\alpha \in (0,1]$, power model is valid on the sphere for some $c_0 \ge \int_0^{\pi} (\theta/a)^{\alpha} \sin\theta d\theta$.

Furthermore, Gneiting (2013) argued that Matérn covariance function is valid on the sphere when the smoothness parameter $\nu \in (0, 1/2)$ and it is not valid if $\nu > 1/2$. Yadrenko (1983) showed that if $K(\cdot)$ is valid isotropic covariance function on $\mathbb{R}^3$ then

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

is a valid isotropic covariance function on the unit sphere (where θ is *gcd*).

4.2 Variogram on a sphere

A random process $X(\cdot)$ on a sphere, Huang et al. (2011) defined, if $E(X(s)) = \mu$ for all $s \in S^2$ and $Var(X(s_1) - X(s_2)) = 2\gamma(\theta_{s_1 s_2})$ and for all $s_1, s_2 \in S^2$ then $X(\cdot)$ is intrinsically statioinary on S^2 where $2\gamma(\cdot)$ is the variogram. The variogram is conditionally negative definite

$$\sum_{i,j=1}^{N} a_i a_j 2\gamma(\theta_{s_1 s_2}) \leq 0, \tag{4.2.1}$$

for any integer N, any constants $a_1, a_2, \ldots, a_N$ with $\sum a_i = 0$, and any locations $s_1, s_2, \ldots, s_N \in S^2$. Immeditely from 4.2.2 for a continuous $2\gamma(\cdot)$ with $\gamma(0) = 0$ the variogram is negative definite if and only if

$$\gamma(\theta) = \sum_{k=0}^{\infty} c_k(1 - P_k(\cos\theta)) \tag{4.2.2}$$

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

In the introduction we pointed out in $\mathbb{R}^d$ one can always construct the variogram if the covariance function is given but not the converse. However, in S^2 Yaglom (1961) argued that for a valid $\gamma(\theta) \quad \theta \in [0, \pi]$ one can always construct covariance function $C(\theta) = c_0 - \gamma(\theta)$ for some $c_0 \geq \int_0^\pi \gamma(\theta)\sin(\theta)d\theta$.

4.3 Random process on a sphere

Jones (1963) showed that a random process on a sphere, can be written as a summation of spherical harmonics $Y_\nu^m(P)$.

A random process $X(P)$ on a unit sphere S^2, where $P = (\lambda, \phi) \in S^2$ where $P = (\lambda, \phi) \in S^2$ with longitude $\lambda \in [-\pi, \pi)$ and latitude $\phi \in [0, \pi]$.. Suppose the porcess is isotropy and continuous in quadratic mean with respect to the location P then the process can be represented by spherical harmonics, $P_\nu^m(\cdot)$ normalized associated Legendre polynomials, with the sum converges in mean square (Li and North (1997); Huang et al. (2012)).

$$X(P) = \sum_{\nu=0}^{\infty} \sum_{m=-\nu}^{\nu} Z_{\nu,m} e^{im\lambda} P_\nu^m(\cos\phi), \tag{4.3.1}$$

Since $\cos(\phi) \in [-1, 1]$ we have $\int_{-1}^{1}[P_\nu^m(cos(\phi))]^2 dcos(\phi) = 1$, and $Z_{\nu,m}$ are complex-valued coefficients satisfying.

$$Z_{\nu,m} = \int_{S^2} X(P) e^{-im\lambda} P_\nu^m(\cos\phi) dP. \tag{4.3.2}$$

Suppose the process $X(P)$ is isotropy with 0 mean (without loss of generality) which implies $E(Z_{\nu,m}) = 0$. Let $P = (\lambda_P, \phi_P)$ and $Q = (\lambda_Q, \phi_Q)$ be two arbitary locations on the sphere, if the covariance function $R(P, Q)$ on S^2 solely depends on the spherical distance between (P, Q) and from 4.2.2, 4.1.6 we can derive the covariance function as follows,

$$
\begin{aligned}
R(P,Q) &= \mathrm{E}(X(P)\overline{X(Q)}) \\
&= \sum_{\nu=0}^{\infty}\sum_{\mu=0}^{\infty}\sum_{m=-\nu}^{\nu}\sum_{n=-\mu}^{\mu} \mathrm{E}(Z_{\nu,m}\overline{Z}_{\mu,n})e^{im\lambda_P}P_\nu^m(\cos\phi_P)e^{-in\lambda_Q}P_\mu^n(\cos\phi_Q) \\
&= \sum_{\nu=0}^{\infty}\frac{(2\nu+1)f_\nu}{2}P_\nu(\cos\theta_{PQ}). \qquad (4.3.3)
\end{aligned}
$$

where $\bar{Z}$ denotes the complex conjugate of Z, θ_{PQ} is the spherical distance, $P_\nu(\cdot)$ is the Legendre polynomial, and $\sum_{\nu=0}^{\infty}(2\nu+1)f_\nu < \infty$. Here, the random variable $Z_{\nu,m}$ satisfies

$$\mathrm{E}(Z_{\nu,m}\overline{Z}_{\mu,n}) = \delta_{\nu,\mu}\delta_{n,m}f_\nu,$$

where $\delta_{a,b} = 1$ if $a = b$, and 0 otherwise.

4.4 Axially symmetry

In the previous sections we discssed why it is necessary to use S^2 instead of R^3 when studying about random processes on Earth and isotropy is often assumed (Yadrenko (1983); Yaglom (1987)). However, many studies have pointed out this assumption is not resonable (Stein (2007); Jun and Stein (2008); Bolin and Lindgren (2011)) for random processes on the sphere primarly on Earth. Stein (2007) argued that Total Ozone Mapping Spectrometer (TOMS) data varies strongly with latitudes and homogeneous models are not suitable. Moreover, aerosol depth (AOD) from Multi-angle Imaging Spectrometer (MISR), Sea Surface Temprature (SST) from RRMM MIcrowave Imager (TMI) are some other example for anisotropy global data on a sphere (on Earth). In order to study non homogenous processes on the sphere Jones (1963) introduces the concept of axially symmetry, where the covarince between two spatial points depend on the longitudes only through their difference between two points.

A rondom process $X(P) : P \in S^2$ on the sphere and let $R(P,Q)$ be a valid covarince finction on the sphere where $P = (L_P, l_P), Q(L_Q, l_Q)$ then $X(P)$ is axially symmetric if and only if

$$R(L_P, L_Q, l_P, l_Q) = R(L_P, L_Q, l_P - l_Q).$$

Currently, to our knowledge there are no methods to test axially symmetry in real data. However, this assumption is more plausible and reasonable when modeling spatial data. For example, temperature, moisture, etc. most likely symmetric on longitudes rather than latitudes. Stein (2007) propose a method to model axially symmetric process on a sphere (the fitted model is not the best, but this was a good start). When modeling spatial data stationary models are less useful; but using the concept of axially symmetry Jun and Stein (2008) proposed a flexible class of parametric covariance models to capture the non-stationarity of global data. Hitczenko and Stein (2012) discussed about the properties of an existing class of models for axially symmetric Gaussian processes on the sphere. They applied first-order

differential operators to an isotropic process. Huang et al. (2012) developed a new representation of axially symmetric process on the sphere and futher introduced some parametric covariance models that are valid on S^2.

if the process is axially symmetric $E(Z_{\nu,m}\overline{Z}_{\mu,n})$ can be expressed as,

$$\mathrm{E}(Z_{\nu,m}\overline{Z}_{\mu,n}) = \delta_{n,m} f_{\nu,\mu,m}.$$

Hence, for an axially symmetric process the covariance function 4.3.3 will be the following form (Huang et al. (2012))

$$\begin{aligned} R(P,Q) &= R(\phi_P, \phi_Q, \lambda_P - \lambda_Q) \\ &= \sum_{m=-\infty}^{\infty} \sum_{\nu=|m|}^{\infty} \sum_{\mu=|m|}^{\infty} f_{\nu,\mu,m} e^{im(\lambda_P - \lambda_Q)} P_\nu^m(\cos\phi_P) P_\mu^m(\cos\phi_Q). \end{aligned} \tag{4.4.1}$$

Inorder to have a valid covariance function, $f_{\nu,\mu,m} = \overline{f}_{\mu,\nu,m}$ and for each fixed integer m, the matrix $F_m(N) = \{f_{\nu,\mu,m}\}_{\nu,\mu=|m|,|m|+1,\ldots,N}$ must be positive definite for all $N \geq |m|$.

$$R(P,Q) = R(\phi_P, \phi_Q, \Delta\lambda) = \sum_{m=-\infty}^{\infty} e^{im\Delta\lambda} C_m(\phi_P, \phi_Q) \quad m = 0, \pm 1, \pm 2, ... \tag{4.4.2}$$

where $\Delta\lambda \in [-\pi, \pi]$ and $\phi_P, \phi_Q \in [0, \pi]$

4.4.1 Properties of $C_m(\phi_P, \phi_Q)$

The covariance function $R(P,Q)$ based on the concept of axially symmetry is clearly defined by both latitudes and longitudes (difference). The following conditions for $C_m(\phi_P, \phi_Q)$ are very important to have a valid covaraince function defined by 4.4.2.

- Hermitian and positive definite.
- $\sum_{m=-\infty}^{\infty} |C_m(\phi_P, \phi_Q)| < \infty$ for $m = 0, \pm 1, \pm 2, ...$
- Is a continuous function.

One can use inverse Fourier transformation to derive C_m based on an axially symmetric covariance function $R(P,Q)$ defined on a sphere, as we have

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

Since $C_m(\phi_P, \phi_Q)$ is continuous and both Hermitian and positive definite, Mercer's theorem (1.1.1) can be directly applied to $C_m(\phi_P, \phi_Q)$ such that there exists an orthonormal basis $\{\psi_{m,\nu}, \nu = 0, 1, \ldots\}$ in L^2 (Huang et al. (2012)) and $C_m(\phi_P, \phi_Q)$ can be given by,

$$C_m(\phi_P, \phi_Q) = \sum_{\nu=0}^{\infty} \eta_{m,\nu} \psi_{m,\nu}(\phi_P) \overline{\psi_{m,\nu}(\phi_Q)},$$

Now the covariance function on a sphere is given by,

$$R(P,Q) = \sum_{m=-\infty}^{\infty} \sum_{\nu=0}^{\infty} \eta_{m,\nu} e^{im\Delta\lambda} \psi_{m,\nu}(\phi_P) \overline{\psi_{m,\nu}(\phi_Q)},$$

Where $\Delta\lambda \in [0, \pi]$, $\eta_{m,\nu} \geq 0$ and $\psi_{m,\nu}(\cdot)$ are the eigen values eigen functions of $C_m(\phi_P, \phi_Q)$ respectively.

In general for covariance function defined on a sphere (Stein (2007)) requires triple summation and required to estimate $O(n^3)$ parameters. In contrast, the covariance function 4.4.2 defined by Huang et al. (2012) requires to estimate $O(n^2)$ parameters which is a huge reduction of comoputational compelextity and we will continue to use this covariance model in our approach on global data generation which is discussed in the next chapter.

Let's consider a real-valued process with a complex valued $C_m(\phi_P, \phi_Q)$ as given below,

$$\begin{aligned} C_m(\phi_P, \phi_Q) &= c_m f(\phi_P, \phi_Q) e^{i\omega_m(\phi_P - \phi_Q)}, \quad c_m \geq 0, \omega_m \in R \\ &= c_m C_m^R(\phi_P, \phi_Q) + i c_m C_m^I(\phi_P, \phi_Q). \end{aligned}$$

Huang et al. (2012) states that if a process is real-valued then the corresponding covariance function $R(P,Q)$ is also real-valued and $C_m(\phi_P, \phi_Q) = \overline{C_{-m}(\phi_P, \phi_Q)}$, The covarince function $R(P,Q)$ on the sphere given by 4.4.2 can be simplified to the following form,

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

There are several covariance models, $R(P,Q)$, valid on a sphere were suggested by Huang et al. (2012) by carefully choosing values for c_m.

Model	c_m	paramters
model 1	: $c_m = Cp^m$	$m = 0, \pm1, \pm2, ... \quad p \in (0,1)$
model 2	: $c_m = \frac{Cp^m}{m}$ and $c_0 = 0$	$m = \pm1, \pm2, ... \quad p \in (0,1)$
model 3	: $c_m = \frac{C}{m^4}$ and $c_0 = 0$	$m = \pm1, \pm2, ...$

Table 4.2: some proposed c_m models

4.5 Longitudinally reversibile process

The idea was first introduced by Stein (2007). Suppose $K(\cdot)$ is a valid covariance function defined on a sphere where,

$$K(L_1, L_2, l_1 - l_2 = K(L_1, L_2, l_2 - l_1) \tag{4.5.1}$$

then underline process is said to be logitudinally reversible. For example the covariance model proposed by Huang et al. (2012) clearly yileds a longitudinally reveresible process as $R(\phi_P, \phi_Q, \Delta\lambda) = R(\phi_P, \phi_Q, -\Delta\lambda)$ and the reversibility holds when $C_{-m}(\phi_P, \phi_Q) = C_m(\phi_P, \phi_Q)$. Now the covariance function reduces to the following,

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

If a random process on the sphere is real valued and longitudinally reversible so is the covariance function, $R(P, Q)$, is real valued then $C_m(\phi_P, \phi_Q)$ is real since $C_{-m}(\phi_P, \phi_Q) = C_m(\phi_P, \phi_Q)$ and $C_m(\phi_P, \phi_Q) = \overline{C_{-m}(\phi_P, \phi_Q)}$.

Chapter 5

Global Data Generation on the Sphere

5.1 Theoretical development

Let $X(P)$ be a complex-valued random process on a unit sphere S^2, where $P = (\lambda, \phi) \in S^2$ with longitude $\lambda \in [-\pi, \pi)$ and latitude $\phi \in [0, \pi]$.

A covariance function for continuous axially symmetric processes on a sphere given by (Huang et al., 2012, proposition 1):

$$R(P,Q) = R(\phi_P, \phi_Q, \Delta\lambda) = \sum_{m=-\infty}^{\infty} e^{im\Delta\lambda} C_m(\phi_P, \phi_Q) \tag{5.1.1}$$

where $\Delta\lambda \in [-\pi, \pi]$, and $C_m(\phi_P, \phi_Q)$ is Hermitian and *p.d.* with $\sum_{-\infty}^{\infty} |C_m(\phi_P, \phi_Q)| < \infty$.

One can derive C_m based on an axially symmetric covariance function $R(P,Q)$ defined on a sphere, as we have

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

since $C_m(\phi_P, \phi_Q)$ is continuous and both Hermitian and positive definite, by Mercer's theorem, there exists an orthonormal basis $\{\psi_{m,\nu}, \nu = 0, 1, \ldots\}$ in L^2, a complex-valued functional Hilbert space on $[o, \pi]$, such that

$$C_m(\phi_P, \phi_Q) = \sum_{\nu=0}^{\infty} \eta_{m,\nu} \psi_{m,\nu}(\phi_P) \overline{\psi_{m,\nu}(\phi_Q)},$$

Where $\eta_{m,\nu} \geq 0$ are the eigen values and $\psi_{m,\nu}(\cdot)$ are the eigen functions.

Further, according to Huang et al. (2012)[remark 2.5] a continuous axially symmetric process, $X(P)$, is given as:

$$X(P) = X(\phi, \lambda) = \sum_{m=-\infty}^{\infty} W_{m,\nu}(\phi) e^{im\lambda} \psi_{m,\nu}(\phi), \tag{5.1.2}$$

where λ is the longitude, ϕ is the latitude and $\psi_{m,\nu}(\cdot)$ is a orthonormal basis. When the process is real and Gaussian, $W_{m,\nu}$ are independent normal random variables. In addition, this process can be viewed as a homogeneous random process on the circle with angular distance given by $\Delta\lambda$. That is, for each ϕ, one can expand $X(P)$ in a Fourier series that is convergent in quadratic mean Roy (1972):

$$X(\phi,\lambda) = \sum_{m=-\infty}^{\infty} W_m(\phi)e^{im\lambda} \tag{5.1.3}$$

,

where

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

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

5.2 Generalization of parametric models

The $R(P,Q)$ given in equation 4.4.2, is clearly a function of both longitude and latitude. The simplest model is the separable model where,

$$R(P,Q) = \tilde{C}(\Delta\lambda)C_m(\phi_P,\phi_Q)$$

In order to make things easier one could assume that $C_m(\phi_P,\phi_Q) = \tilde{C}_m(\phi_P-\phi_Q)$ only depends on the difference of ϕ_P and ϕ_Q, Huang et al. (2011) proposed a simple separable covariance function when both covariance components are exponential

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

Lets assume,

$$C_m(\phi_P,\phi_Q) = c_m e^{-a_m|\phi_P-\phi_Q|}(\cos\omega_m(\phi_P-\phi_Q)+i\sin\omega_m(\phi_P-\phi_Q)), \quad C_m\geq 0, a_m\geq 0, \omega_m\in R.$$

(Huang et al., 2012, Remark 2.4)states that for real-valued process, the covariance function $R(P,Q)$ is also real-valued, and

$$R(P,Q) = C_{0,R}(\phi_P,\phi_Q) + 2\sum_{m=1}^{\infty} cos(m\Delta\lambda)C_{m,R}(\phi_P,\phi_Q) - sin(m\Delta\lambda)C_{m,I}(\phi_P,\phi_Q),$$

where $C_m(\phi_P,\phi_Q)$=$C_{m,R}(\phi_P,\phi_Q)$+$iC_{m.I}(\phi_P,\phi_Q)$.
$C_{0,R}(\phi_P,\phi_Q)$=$c_0e^{-a_0|\phi_P-\phi_Q|}\cos\omega_0(\phi_P-\phi_Q)$, $C_{m,R}(\phi_P,\phi_Q)$=$c_me^{-a_m|\phi_P-\phi_Q|}\cos\omega_m(\phi_P-\phi_Q)$,
$C_{m,I}(\phi_P,\phi_Q)$=$c_me^{-a_m|\phi_P-\phi_Q|}\sin\omega_m(\phi_P-\phi_Q)$

If one takes $a_m = a$, $\omega_m = mu$ we can get the following form for $R(P,Q)$,

$$R(P,Q) = c_0e^{-a|\phi_P-\phi_Q|} + 2e^{-a|\phi_P-\phi_Q|}\sum_{m=1}^{\infty} c_m\cos[m\theta(P,Q,u)],$$

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

Moreover, by carefully choosing functions for $C_m(\phi_P, \phi_Q)$ Huang et al. (2012) proposed some nonseparable covariance models ($R(P,Q)$) models valid on the sphere,

$$R(P,Q) = Ce^{-a|\phi_P-\phi_Q|}\frac{1-p^2}{1-2p\cos\theta(P,Q,u)+p^2} \tag{5.2.1}$$

$$R(P,Q) = Ce^{-a|\phi_P-\phi_Q|}\log\frac{1}{(1-2p\cos\theta(P,Q,u)+p^2)} \tag{5.2.2}$$

$$R(P,Q) = 2Ce^{-a|\phi_P-\phi_Q|}\left(\frac{\pi^4}{90} - \frac{\pi^2\theta^2(P,Q,u)}{12} + \frac{\pi\theta^3(P,Q,u)}{12} - \frac{\theta^4(P,Q,u)}{48}\right) \tag{5.2.3}$$

There is one big disadvantage for all of them. They are assumed not only stationarity on longitudes, but stationarity on latitudes as well.

Modifying the covariance models to include non-stationarity

1. We have noticed that when $\phi_P = \phi_Q$, the first model reduces to

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

and if $\Delta\lambda = 0$, the variance over all latitudes would be constant. This is not supposed to be the case, since both MSU data and TOMS data in figures **??** and **??** shows that variance is highly depending on the latitude.

2. A modification of the above approach is to replace the function

$$C(\phi_P, \phi_Q) = Ce^{-a|\phi_P-\phi_Q|}$$

by a non-stationary covariance function, which depends on the latitudes, even when $\phi_P = \phi_Q$. Here are the two functions that have been used in our analysis.

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

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

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

$$\tilde{C}(\phi_P,\phi_P) = C_1(C_2 - 2e^{-a|\phi_P|} + 1),$$

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

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

   Proposition 5.2.1 *A more general non stationary covariance function is given as following. If $C(\cdot) = C(x-y)$ is the stationary covariance function and $f(\omega) \geq 0$ is the corresponding spectral density, then*

   $$\tilde{C}(x,y) = C_2 - C(x) - C(y) + C(x-y), \qquad C_2 \geq \int_{-\infty}^{\infty} dF(\omega) = \int_{-\infty}^{\infty} f(\omega)d\omega > 0$$

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

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

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

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

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

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

So we propose a five-parameter model for the covariance on a sphere

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

where $C_1 > 0, C_2 > 0, a, u, p$ can be estimated from the data.

5.3 Method development

We can construct normal independent (complex) random variate $W_m(\phi)$ associated with the variance-covariance matrix $C_m(\phi_P, \phi_Q)$ to construct an axially symmetric process. Then

finite summation can be used to approximate above (5.1.3) infinite summation as given below,

$$X(P) = X(\phi, \lambda) = \sum_{m=-N}^{N} W_m(\phi) e^{im\lambda} \tag{5.3.1}$$

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

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

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

$$C_{-m}(\phi_P, \phi_Q) = \overline{C_m(\phi_P, \phi_Q)}, \quad \text{for } m = 1, 2, \cdots, N \tag{5.3.2}$$

Lets write $W_m(\phi) = W_m^r(\phi) + iW_m^i(\phi)$ in terms of a real component and an imaginary component. We also write $C_m(\phi_P, \phi_Q) = C_m^r(\phi_P, \phi_Q) + iC_m^i(\phi_P, \phi_Q)$ and with the relationship 5.3.2 above, we have

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

Now,

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

If we let $W_{-m}(\phi) = \overline{W_m(\phi)}$, then the covariance function would satisfy the above relationship 5.3.2. In addition, we will set the following,

$$Cov(W_m^r(\phi_P), W_m^r(\phi_Q)) = Cov(W_m^i(\phi_P), W_m^i(\phi_Q)) = \frac{1}{2} C_m^r(\phi_P, \phi_Q), \tag{5.3.3}$$

$$Cov(W_m^i(\phi_P), W_m^r(\phi_Q)) = -Cov(W_m^r(\phi_P), W_m^i(\phi_Q)) = \frac{1}{2} C_m^i(\phi_P, \phi_Q). \tag{5.3.4}$$

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

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

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

The process given by (5.1.3) is now simplified as the following (real) process,

$$\begin{aligned} X(P) &= \sum_{m=-N}^{N} W_m(\phi)e^{im\lambda} = W_0(\phi) + \sum_{m=1}^{N} W_m(\phi)e^{im\lambda} + \sum_{m=-1}^{-N} W_m(\phi)e^{im\lambda} \\ &= W_0(\phi) + \sum_{m=1}^{N} W_m(\phi)e^{im\lambda} + \sum_{m=1}^{N} \overline{W_m(\phi)}e^{-im\lambda} \\ &= W_0(\phi) + \sum_{m=1}^{N} \big[(W_m^r(\phi) + iW_m^i(\phi))(\cos(m\lambda) + i\sin(m\lambda)) \\ &\quad + (W_m^r(\phi) - iW_m^i(\phi))(\cos(m\lambda) - i\sin(m\lambda))\big] \\ &= W_0(\phi) + 2\sum_{m=1}^{N} \big[W_m^r(\phi)\cos(m\lambda) - W_m^i(\phi)\sin(m\lambda)\big]. \end{aligned} \tag{5.3.5}$$

5.3.1 Data generation

Now for each fixed $m = 0, 1, 2, \cdots, N$, we consider $W_m(\phi) = W_m^r(\phi) + iW_m^i(\phi)$ then $W_m^*(\phi) = W_m^r(\phi) - iW_m^i(\phi)$ (where $W_m^*(\phi)$ is the conjugate of $W_m(\phi)$). We may assume that $W_m^r(\phi)$ and $W_m^i(\phi)$ are independent, each following a (normal) distribution with mean zero and the same variance $\sigma_m^2(\phi) = \frac{1}{2}C_m^r(\phi,\phi)$, ($C_m^i(\phi,\phi)=0$ implies $W_m^r(\phi)$ and $W_m^i(\phi)$ are uncorrelated, or independent for Gaussian). From 1.1.2, $W_m(\phi)$ is the circularly-symmetric complex random variable (Gallager (2008)).

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

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

Now we calculate the covariance matrix $K_W = E(\underset{\sim}{W}_m\underset{\sim}{W}_m^*)$ (where $\underset{\sim}{W}_m^*$ is the conjugated transpose) and pseudo-covariance $M_W = E(\underset{\sim}{W}_m\underset{\sim}{W}_m^T)$. Further, from 1.1.2 a complex random vector is circularly-symmetric if and only if M_W is zero.

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

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

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

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

$$
\begin{aligned}
K_W &= E(\underline{W}_m \underline{W}_m^*) \\
&= \begin{pmatrix} E[W_m(\phi_1)W_m^*(\phi_1)] & E[W_m(\phi_1)W_m^*(\phi_2)] & \cdots & E[W_m(\phi_1)W_m^*(\phi_{n_l})] \\ E[W_m(\phi_2)W_m^*(\phi_1)] & E[W_m(\phi_2)W_m^*(\phi_2)] & \cdots & E[W_m(\phi_2)W_m^*(\phi_{n_l})] \\ \vdots & \vdots & \ddots & \vdots \\ E[W_m(\phi_{n_l})W_m^*(\phi_1)] & E[W_m(\phi_{n_l})W_m^*(\phi_2)] & \cdots & E[W_m(\phi_{n_l})W_m^*(\phi_{n_l})] \end{pmatrix} \\
&= \begin{pmatrix} C_m^r(\phi_1, \phi_1) & C_m^r(\phi_1, \phi_2) + iC_m^i(\phi_1, \phi_2) & \cdots & C_m^r(\phi_1, \phi_{n_l}) + iC_m^i(\phi_1, \phi_{n_l}) \\ C_m^r(\phi_2, \phi_1) - iC_m^i(\phi_2, \phi_1) & C_m^r(\phi_2, \phi_2) & \cdots & C_m^r(\phi_2, \phi_{n_l}) + iC_m^r(\phi_2, \phi_{n_l}) \\ \vdots & \vdots & \ddots & \vdots \\ C_m^r(\phi_{n_l}, \phi_1) - iC_m^i(\phi_{n_l}, \phi_1) & C_m^r(\phi_{n_l}, \phi_2) - iC_m^r(\phi_{n_l}, \phi_2) & \cdots & C_m^r(\phi_{n_l}, \phi_{n_l}) \end{pmatrix} \\
&= \begin{pmatrix} C_m^r(\phi_1, \phi_1) & C_m^r(\phi_1, \phi_2) & \cdots & C_m^r(\phi_1, \phi_{n_l}) \\ C_m^r(\phi_2, \phi_1) & C_m^r(\phi_2, \phi_2) & \cdots & C_m^r(\phi_2, \phi_{n_l}) \\ \vdots & \vdots & \ddots & \vdots \\ C_m^r(\phi_{n_l}, \phi_1) & C_m^r(\phi_{n_l}, \phi_2) & \cdots & C_m^r(\phi_{n_l}, \phi_{n_l}) \end{pmatrix} \\
&\quad + i \begin{pmatrix} 0 & C_m^i(\phi_1, \phi_2) & \cdots & C_m^i(\phi_1, \phi_{n_l}) \\ -C_m^i(\phi_2, \phi_1) & 0 & \cdots & C_m^i(\phi_2, \phi_{n_l}) \\ \vdots & \vdots & \ddots & \vdots \\ -C_m^i(\phi_{n_l}, \phi_1) & -C_m^i(\phi_{n_l}, \phi_2) & \cdots & 0 \end{pmatrix} \\
&= Re(K_W) + iIm(K_W)
\end{aligned}
$$

Now,

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

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

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

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

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

Since $\underset{\sim}{W}_m$ is circularly-symmetric from 1.1.11 we can get the following results,

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

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

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

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

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

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

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

where $Z = \{z_1, z_2, \ldots, z_{n_l}, z_1^*, z_2^*, \ldots, z_{n_l}^*\}$ and each $z_i \sim N(0,1)$ hence we can get $\underset{\sim}{W}_m$. Now for each latitude $\phi_l, l = 1, 2, \cdots, n_l$ and $\lambda_k, k = 1, 2, \cdots, n_L$ ($N = n_L/2$), we denote the axially symmetric data (real) as $X(\phi_l, \lambda_k)$. These random variates can be obtained from the equation (5.3.5), let's rewrite the equation as follows,

$$X(\phi_l, \lambda_k) = W_0(\phi_l) + 2\sum_{m=1}^{N} \left[W_m^r(\phi_l)\cos(m\lambda_k) - W_m^i(\phi_l)\sin(m\lambda_k)\right] \qquad (5.3.6)$$

5.3.2 Pseudo-code

- Choose a cross covariance function, $R(P, Q)$
- Initialize the parameters (C_1, C_2, a, u, p) and choose a resolution $\phi_1, \ldots, \phi_{n_l}, \lambda_1, \ldots, \lambda_{n_L}$ (or $n_l \times n_L$),
- Derive $C_m(\phi_P, \phi_Q)$ based on $R(P, Q)$ where $m = 0, 1, \ldots, n_L/2$,
  1. for each m get $Re(K_W)$ and $Im(K_W)$ hence obtain K_V
  2. use SVD to get $\underset{\sim}{V}_m$ $(n_l - tuples)$
  3. get $\underset{\sim}{W}_m$'s from $\underset{\sim}{V}_m$
- apply the equation (5.3.6) to generate grid data.

5.4 Property of MOM

we might have to omit this section, add a discussion about variogram

We should conform the validity of proposed covariance functions $R(P, Q)$. Since $R(P, Q)$ functions are cross covariance functions, we can compute the empirical covariance, by method of moments and compare.

We can estimate the cross covariance between any two arbitrary latitudes at each longitudinal difference (empirical covariance) based on method of moments and compare it with the theoretical $R(P, Q)$ values. According to Wackernagel (2013) a cross covariance function is not an even function so does the proposed $R(P, Q)$ functions, *i.e.* $R(P, Q, \Delta\lambda) \neq R(P, Q, -\Delta\lambda)$. The empirical (cross) covariance between any two latitudes ϕ_P and ϕ_Q with a zero mean processes can be given as follows,

$$\hat{R}(\phi_P, \phi_Q, \theta) = \frac{1}{n_L} \sum_{k=1}^{n_L} (X(\phi_P, \theta + \lambda_k) \cdot X(\phi_l, \lambda_k)), \tag{5.4.1}$$

where $\theta = m\Delta\lambda$.

In general, the spatial processes are stationary at a given latitude but not within latitudes. If the processes are not zero mean once could simply substract the product of means from the above MOM estimator. The above estimate is clearly unbiased as we have, for fixed two locations $P, Q \in S^2$,

$$E[\hat{R}(\phi_P, \phi_Q, \theta)] = R(P, Q).$$

Later we will prove this estimator is consistent, *i.e.,* for any $\varepsilon > 0$,

$$\lim_{n \to \infty} P(|\hat{R}(\phi_P, \phi_Q, \theta) - R(P, Q)| > \varepsilon) = 0.$$

5.4.1 Results

Simulated data sample:

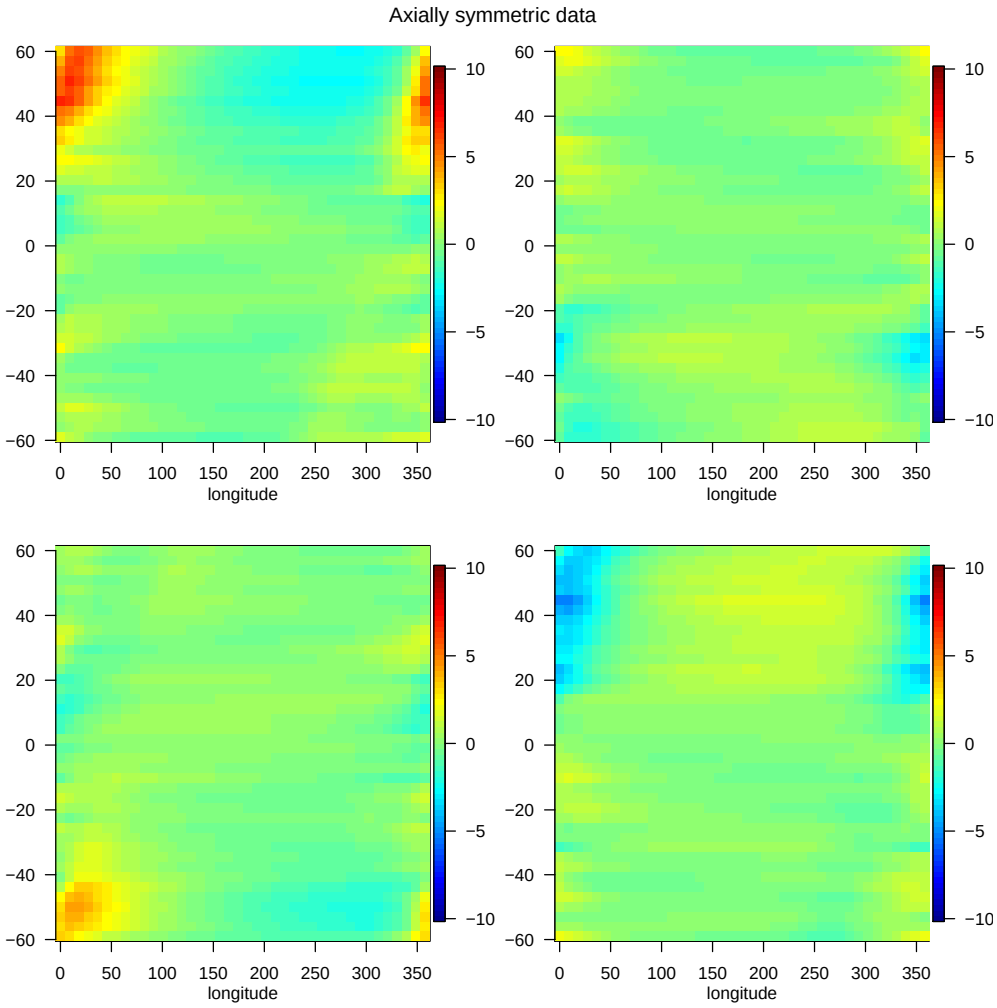

Figure 5.1: Four consecutive axially symmetric data snapshots based on model 2, grid resolution $2^0 \times 1^0$ (data scale -10 and 10).

Comprison of the proposed models with MOM estimates:

- Model 1

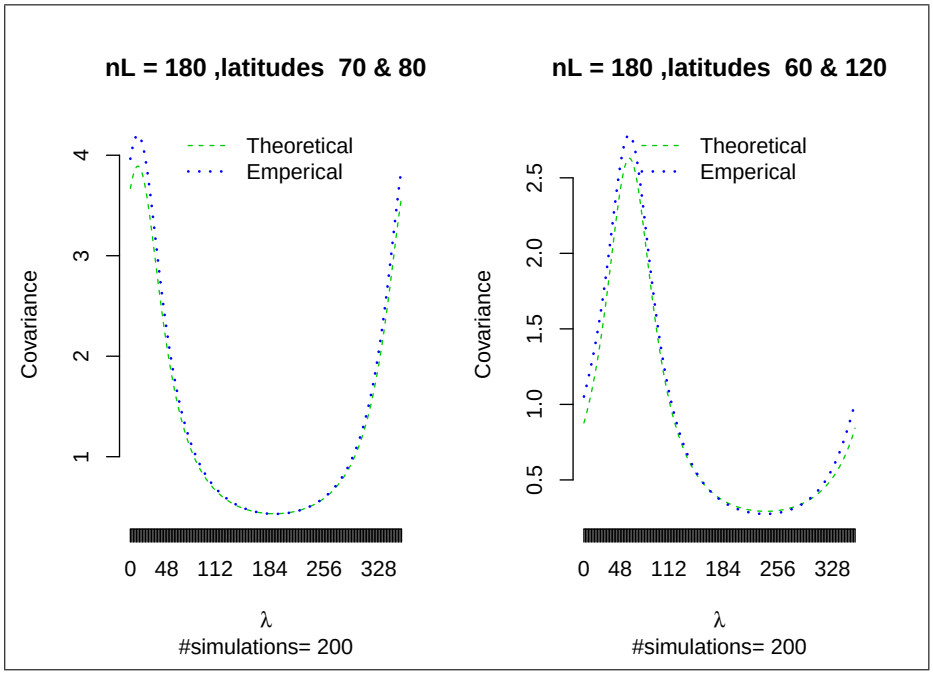

Figure 5.2: Cross covariance comparison of model1

- Model 2

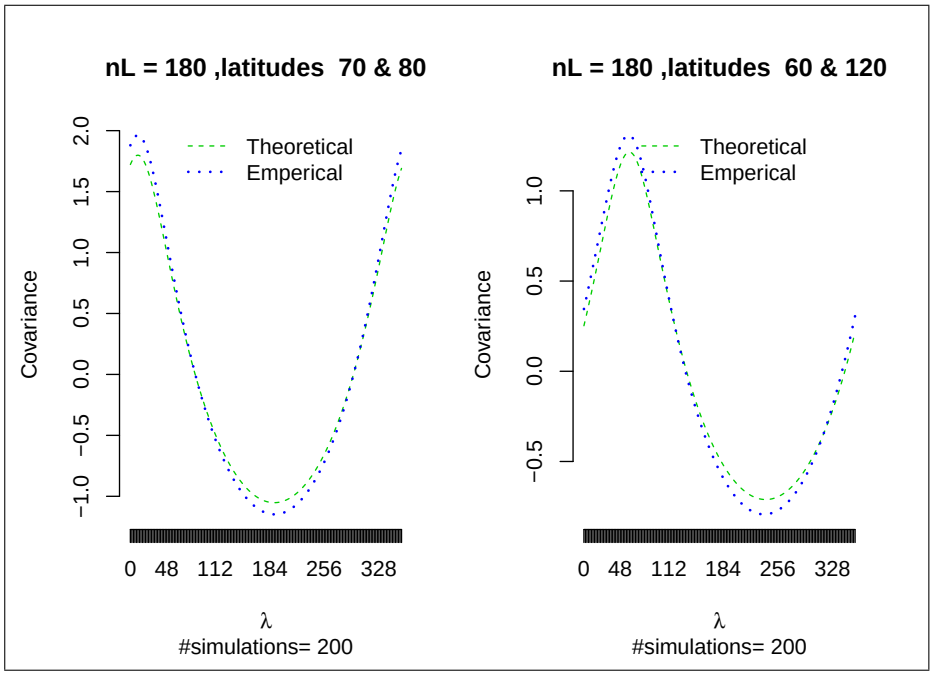

Figure 5.3: Cross covariance comparison of model2

- Model 3

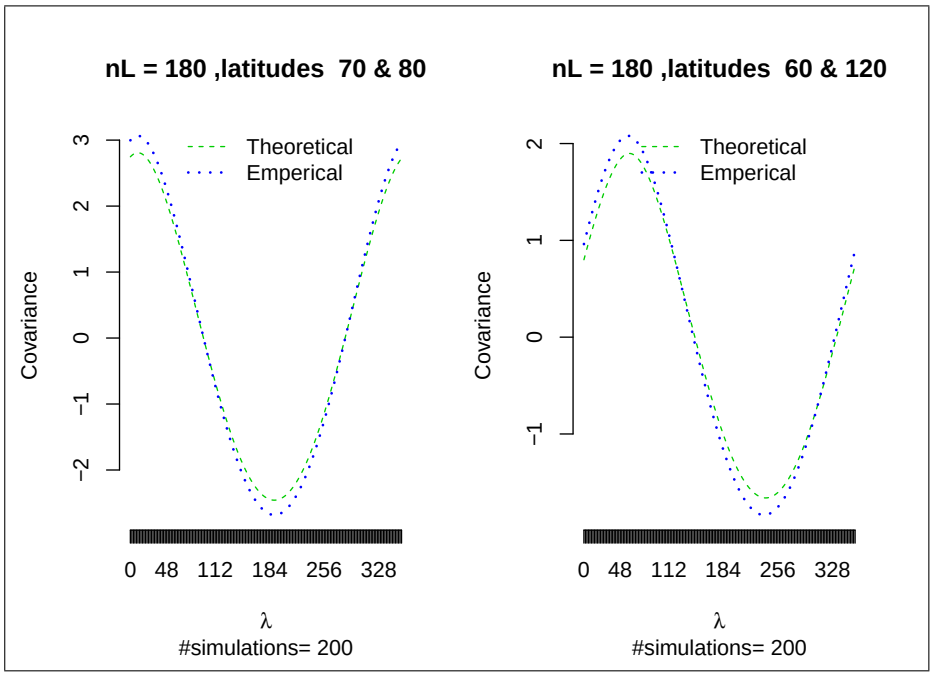

Figure 5.4: Cross covariance comparison of model3

Chapter 6

Future Research (due August 28)

Future research work !!

Bibliography

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.

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.

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.

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

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

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.

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

Roy, R. (1972). Spectral analysis for a random process on the circle. *Journal of Applied Probability*, pages 745–757.

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

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. (1987). *Correlation Theory of Stationary and Related Random Functions: Vol.: 1: Basic Results*. Springer-Verlag.

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