Space-time Covariance Function on Sphere

Xinyue Chang, Haoyan Hu, Xingche Guo

December 3, 2018

Overview

1. Introduction & Background
2. Global Space-time Models for Climate Ensembles
   - Mothodology
   - Computation
   - Application
3. Evolutionary Spectrum Approach to Incorporate Descriptors on Global Processes

Introduction & Background

Definitions and Background

$Z(L, l, t)$ is a spatio-temporal process on $\mathcal{S}_2 = \{x \in \mathbb{R}^3 : ||x|| = R\}$, radius $R \in (0, +\infty)$, $L \in [-90, 90]$ and $l \in [-180, 180]$.

- $\theta = \arccos\{\sin(L_1)\sin(L_2) + \cos(L_1)\cos(L_2)\cos(l_1 - l_2)\}$.
  Geodesic distance $= R\theta$, chordal distance $= 2R\sin(\theta/2)$.
- Smoothness of Processes on a Sphere.

Covariance Models on Sphere:

- Validity
- Isotropy
- Axially sysmmetry

Approaches to Modeling on Sphere

Axially symmetric: $\mathrm{cov}\{Z(L_1, l_1), Z(L_2, l_2)\} = K(L_1, L_2, l_1 - l_2)$
Longitudinal reversible: $K(L_1, L_2, l_1 - l_2) = K(L_1, L_2, l_2 - l_1)$

- Differential Operator
- Spherical Harmonic Representation
- Stochastic Partial Differential Equations
- Kernel Convolution
- Deformations
- Multi-Step Spectrum

Global Space-time Models for Climate Ensembles

Global Space-time Model for Climate Ensembles

We denote with L_m for $m = 1, \cdots, M$ the latitude, with l_n for $n = 1, \cdots, N$ the longitude, with $t = 1, \cdots, T$ the time. $\mathbf{T}_r$ is a vector ordered by time, longitude, latitude from rth realization.

$$\text{Overall Model:} \quad \mathbf{T}_r = \boldsymbol{\mu} + \boldsymbol{\varepsilon}_r, \quad \boldsymbol{\varepsilon}_r \sim \mathcal{N}(\mathbf{0}, \Sigma(\boldsymbol{\theta})), \quad r = 1, \cdots, R$$

where $\boldsymbol{\varepsilon}_r \overset{d}{=} (\tilde{\varepsilon}_1, \cdots, \tilde{\varepsilon}_T)$.

$$\text{Temporal Structure:} \quad \tilde{\varepsilon}_t = \boldsymbol{\Phi}\tilde{\varepsilon}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathcal{N}(\mathbf{0}, \Sigma_s), \quad \tilde{\varepsilon}_1 \sim \mathcal{N}(\mathbf{0}, \Sigma_s)$$

where $\boldsymbol{\Phi} = \mathrm{diag}\left(\varphi_{Q(L_1,\ell_1)}, \ldots, \varphi_{Q(L_M,\ell_N)}\right)$, where $Q(L, l) = 1$ if land fraction of pixel (L, l) is greater than 50% and $Q(L, l) = 0$ otherwise.

Spectral Modeling for A Single Latitude Band

Assume axially symmetry, K_L(constructing Σ_s) is only a function of the longitudinal lag $l_n = 2\pi n/N$, $n = 0, \cdots, N-1$ and is symmetric about π. A modified Matern density:

$$f_L(c) = \frac{\phi_L}{\left(\alpha_L^2 + 4\sin^2(c/N\pi)\right)^{v_L+1/2}}, \quad c = 0, \cdots, N-1$$

Thus,

$$\begin{aligned} K_L(l_n) &= \frac{1}{N}\sum_{c=0}^{N-1} f_L(c) exp(icl_n) \\ &= \frac{1}{N}\sum_{c=0}^{N-1} f_L(c)\cos(cl_n) \end{aligned}$$

So Σ_s is circulant and symmetric.

Spectral Modeling for Multiple Latitude Bands

Assume longitudinal reversible. Specify coherence as

$$\rho_{L_m,L_{m'}}(c) = \left[\frac{\xi}{\left\{1+4\sin^2(c/N\pi)\right\}^{\tau}}\right]^{|L_m-L_{m'}|}$$

and null phase of $f_{L_m,L_{m'}}(c)$. So

$$f_{L_m,L_{m'}}(c) = \rho_{L_m,L_{m'}}(c)\sqrt{f_{L_m}(c)f_{L'_m}(c)}.$$

And cross-covariance function

$$\begin{aligned} K_{L_m,L'_m}(l_n) &= \frac{1}{N}\sum_{c=0}^{N-1} f_{L_m,L_{m'}}(c) exp(icl_n) \\ &= \frac{1}{N}\sum_{c=0}^{N-1} f_{L_m,L'_m}(c)\cos(cl_n) \end{aligned}$$

So cross-covariance (off-diagonal of Σ_s) is circulant and symmetric.

REML

Define $\boldsymbol{D}_r = \mathbf{T}_r - \bar{\mathbf{T}}, r = 1, \cdots, R$ and $\boldsymbol{D} = (\boldsymbol{D}_1, \cdots, \boldsymbol{D}_R)$.

$$l(\boldsymbol{\theta}; \mathbf{D}) = -\frac{TNM(R-1)}{2}\log(2\pi) - \frac{1}{2}(R-1)\log(\det(\Sigma(\boldsymbol{\theta}))) - \frac{1}{2}TNM\log(R) - \frac{1}{2}\sum_{r=1}^{R}\mathbf{D}_r^T(\Sigma(\boldsymbol{\theta}))^{-1}\mathbf{D}_r$$

Maximize $l(\boldsymbol{\theta}; \mathbf{D})$ with constraints $\phi_L > 0, \alpha_L > 0, \nu_L > 0$ for single bands and $\xi \in (0,1), \xi < 5^{\tau}, \varphi_0 > 0, \varphi_1 > 0$ for multiple bands.

The problem reduces to:

- How to compute Σ_s fast?
- How to compute Σ^{-1} and $\det(\Sigma)$ fast?

Fast Computation for Σ_s (Single band case)

- Recall: Σ_s is the covariance matrix for a single latitude band, which is a symmetric circulant matrix.
- Complex eigenvector matrix $\boldsymbol{U}$ of a circulant matrix is the same for all circulant matrices.
- Eigenvalues of a circulant matrix can be computed through fast Fourier transform (FFT).
- So the covariance matrix is diagonal after discrete Fourier transform. It's sufficient to compute the first row of Σ_s.

Fast Computation for Σ^{-1}

- Structure of Σ:

$$\Sigma = \Big(\mathrm{cov}(\tilde{\varepsilon}_s, \tilde{\varepsilon}_t)\Big)_{1 \le s,t \le T}$$

where:

$$\mathrm{var}(\tilde{\varepsilon}_t) = \sum_{i=0}^{t-1} \mathbf{\Phi}^i \Sigma_s \mathbf{\Phi}^i, \qquad \mathrm{cov}(\tilde{\varepsilon}_s, \tilde{\varepsilon}_t) = \sum_{i=0}^{s-1} \mathbf{\Phi}^i \Sigma_s \mathbf{\Phi}^{i+t-s}, \textit{ for } s < t$$

Fast Computation for Σ^{-1}

Recall simple non-stationary AR(1) model:

$$u_t = \rho u_{t-1} + e_t, \qquad 2 \le t \le T$$
$$e_t \sim \mathcal{N}(0, \sigma^2), \quad u_1 \sim \mathcal{N}(0, \sigma^2)$$

Then $\Sigma = \Big(\mathrm{cov}(u_s, u_t)\Big)_{1 \le s,t \le T}$, where: $\mathrm{var}(u_t) = \Big(\sum_{i=0}^{t-1} \rho^{2i}\Big)\sigma^2$, $\mathrm{cov}(u_s, u_t) = \mathrm{var}(u_t)\rho^{t-s}, s < t$.

$$\Sigma^{-1} = \frac{1}{\sigma^2}\begin{pmatrix} 1+\rho^2 & -\rho & 0 & \dots & \dots & \dots & 0 \\ -\rho & 1+\rho^2 & -\rho & \ddots & & & \vdots \\ 0 & -\rho & 1+\rho^2 & -\rho & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & -\rho & 1+\rho^2 & -\rho & 0 \\ \vdots & & & \ddots & -\rho & 1+\rho^2 & -\rho \\ 0 & \dots & \dots & \dots & 0 & -\rho & 1 \end{pmatrix}$$

Fast Computation for Σ^{-1}

Similarly, for vectorized non-stationary AR(1) model:

$$
\Sigma^{-1} = \begin{pmatrix}
A & B & O & \dots & \dots & \dots & O \\
B^T & A & B & \ddots & & & \vdots \\
O & B^T & \ddots & \ddots & \ddots & & \vdots \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
\vdots & & \ddots & \ddots & \ddots & \ddots & O \\
\vdots & & & \ddots & \ddots & A & B \\
O & \dots & \dots & \dots & O & B^T & C
\end{pmatrix}
$$

Where:

$$
A = \Sigma_s^{-1} + \mathbf{\Phi}\Sigma_s^{-1}\mathbf{\Phi}, \quad B = -\mathbf{\Phi}\Sigma_s^{-1}, \quad C = \Sigma_s^{-1}
$$

Hence, in order to compute Σ^{-1}, we only need to compute A, B, an C.

Fast Computation for $\det(\Sigma)$

$$\det\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} = \det(S_{11} - S_{12}S_{22}^{-1}S_{21})\det(S_{22})$$

Define $\Sigma_{T\times T}^{-1} := S^{(T)}$, then

$$S^{(T)} = \begin{pmatrix} A & \tilde{\boldsymbol{B}} \\ \tilde{\boldsymbol{B}}^T & S^{(T-1)} \end{pmatrix}$$

Where $\Sigma_{(T-1)\times(T-1)}^{-1} := S^{(T-1)}$, $\tilde{\boldsymbol{B}} = (B, O, \ldots, O)$.
Compute $\det(\Sigma)$ iteratively:

$$\begin{aligned} \det(S^{(T)}) &= \det(A - \tilde{\boldsymbol{B}} S^{(T-1)} \tilde{\boldsymbol{B}}')\det(S^{(T-1)}) \\ &= det(\Sigma_s^{-1})\det(S^{(T-1)}) = \cdots = \left(det(\Sigma_s^{-1})\right)^T \end{aligned}$$

Hence $\det(\Sigma) = \left(det(\Sigma_s)\right)^T$.

Fast Computation for Multiple Bands

- Σ_s is the covariance matrix for multiple latitude bands.
- Σ_s is a block circulant matrix.
- We can still apply the fast computation methods as for the single band, however, the speed would be slow since the computation of Σ_s^{-1} is slow.
- Can we take advantage of the block circulant structure to accelerate our algorithm?

FT to the REML

Applying DFT to both $\boldsymbol{D}_r$ and $\Sigma = \boldsymbol{H}(\Sigma_s, \boldsymbol{\Phi})$:

$$\boldsymbol{Z}_r = (I_T \otimes \boldsymbol{U}^*)\boldsymbol{D}_r, \quad \Lambda = \boldsymbol{H}(\Lambda_s, \boldsymbol{\Phi}), \ \textit{where } \Lambda_s = \boldsymbol{U}^*\Sigma_s\boldsymbol{U}$$

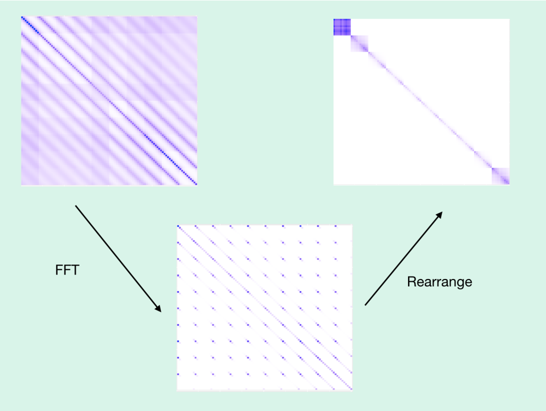
Then the REML becomes:

$$\begin{aligned} l^*(\boldsymbol{\theta}; \mathbf{Z}) = & -\frac{TNM(R-1)}{2}\log(2\pi) - \frac{1}{2}(R-1)\log(\det(\Lambda(\boldsymbol{\theta}))) \\ & - \frac{1}{2}TNM\log(R) - \frac{1}{2}\sum_{r=1}^{R} Re(\mathbf{Z}_r^T(\Lambda(\boldsymbol{\theta}))^{-1}\mathbf{Z}_r) \end{aligned}$$

Rearrange Λ_s and $\boldsymbol{Z}_r$

$$\Sigma_s \xrightarrow{FFT} \Lambda_s \xrightarrow{Rearrange} \tilde{\Lambda}_s$$

$$\boldsymbol{D}_r \xrightarrow{FFT} \boldsymbol{Z}_r \xrightarrow{Rearrange} \tilde{\boldsymbol{Z}}_r$$



Summary of algorithm

- **Single band**: Estimate ϕ_L, α_L and ν_L for each latitude band by maximizing single band REML
  - Step1: Initialize ϕ_L, α_L, ν_L and $\mathbf{\Phi}$.
  - Step2: Compute Σ_s by FFT, apply FFT to data on longitude.
  - Step3: Compute Σ^{-1} and $\det(\Sigma)$.
  - Step4: Use R function "constrOptim" to find the maximum iteratively.
- **Multiple bands**: Estimate ξ, τ and $\mathbf{\Phi}$ using data from multiple bands and treat ϕ_L, α_L and ν_L as known.
  - Step1: Initialize ξ, τ, and $\mathbf{\Phi}$.
  - Step2: Apply FFT and rearrange to Σ_s, apply FFT to data on longitude and reorder.
  - Step3: Compute $\tilde{\Lambda}^{-1}$ and $\det(\tilde{\Lambda})$.
  - Step4: Use R function "constrOptim" to find the maximum iteratively.

Practical Comments

- Fast Fourier transformed data and actually changed the original REML equation.
- But if $\phi_0 = \phi_1$, the reml is the same after FFT.

Dataset: collection

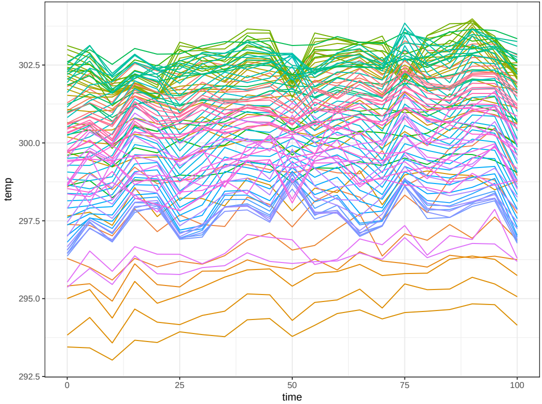
CO_2 forcing increases linearly to 700 ppm over 200 years beginning in 2010 and stays constant thereafter: Atm model, 2000 to 2100 every 5 years ($T = 21$), five realizations ($R = 5$), 42×96 grids ($M = 42, N = 96$).

Dataset: preprocessing

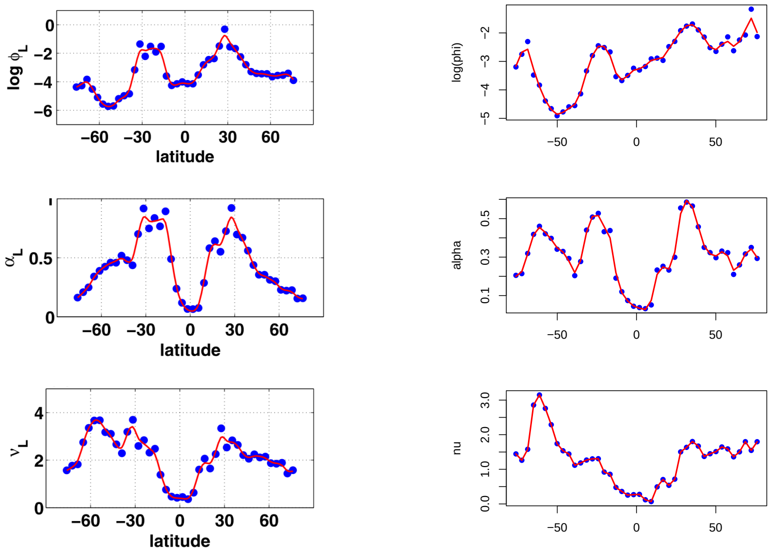
Yearly average temperature at surface: $\mathbf{T}_r$

Land fraction: obtained from January in 2010 of the first run, assuming unchange over time.

A plot of temprature trend at all locations of latitude = -1.86 band at the first run.

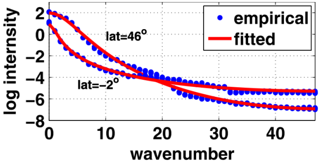


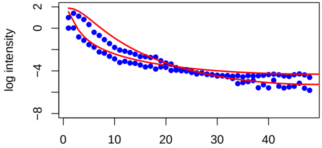
Results for Single Bands



Results

Empirical and fitted log periodogram for two different latitudinal bands.





Multiple bands:

	ξ	τ	φ_0	φ_1
Results in paper	0.9696	0.2080	0.1141	0.1010
Our Results	0.9553	0.2573	1.3×10^{-5}	0.005538

Practical Comments

- Single: 0.029 vs 17.039, Multiple: 1.282 vs 218.815.
- Setting $\phi_0 = \phi_1$ does not affect the single band results too much.
- Sampling data can help increase computation speed and wouldn't cause large estimation error. But lag of 5 years should be limit. We tried data that sampled every 20 years, which did not work well.
- The method heavily depends on the data structure: grid data, fully observed on longitudinal circle and axially symmetry. For some real global datasets, the modified spectral density and FFT are not valid.

Evolutionary Spectrum Approach to Incorporate Descriptors on Global Processes

Temporal Part:

Use AR(2) model with different coefficients for every grid point instead of AR(1) model with only different coefficients for land-ocean.

$$\text{Previous:} \quad \tilde{\varepsilon}_t = \mathbf{\Phi}\tilde{\varepsilon}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim \mathcal{N}\left(\mathbf{0}, \Sigma_s\right), \quad \tilde{\varepsilon}_1 \sim \mathcal{N}\left(\mathbf{0}, \Sigma_s\right)$$

where $\varepsilon_r \stackrel{d}{=} (\tilde{\varepsilon}_1, \cdots, \tilde{\varepsilon}_T)$, $\mathbf{\Phi} = \mathrm{diag}\left(\varphi_{Q(L_1,\ell_1)}, \ldots, \varphi_{Q(L_M,\ell_N)}\right)$

$$\text{Here:} \quad \tilde{\varepsilon}\left(t_k; r\right) = \mathbf{\Phi_1}\tilde{\varepsilon}\left(t_k - 1; r\right) + \mathbf{\Phi_2}\tilde{\varepsilon}\left(t_k - 2; r\right) + \boldsymbol{SH}\left(t_k; r\right)$$

where $\mathbf{\Phi_1}$ and $\mathbf{\Phi_2}$ are two $\boldsymbol{NM} X \boldsymbol{NM}$ diagonal matrices with the AR coefficients for each location and $\boldsymbol{S}$ is an $\boldsymbol{NM} X \boldsymbol{NM}$ diagonal matrix with the associated standard deviation.

$H_r(L_m, l_n, t_k)$ are independent across time and describe the spatial dependence across the sphere. Model the process in the spectral domain across longitudes, and then to model the dependence across latitudes:

$$\mathbf{H}_r\left(L_m, l_n, t_k\right) = \sum_{c=0}^{N-1} f_{L_m, l_n}(c) \exp\left(i l_n c\right) \tilde{\mathbf{H}}_r\left(c, L_m, t_k\right)$$

$$\operatorname{corr}\left\{\tilde{\mathbf{H}}_r\left(c, L_m, t_k\right), \tilde{\mathbf{H}}_{r'}\left(c', L_{m'}, t_k'\right)\right\} = \mathbf{1}\left\{c = c', k = k', r = r'\right\} \rho_{L_m, L_m(c)}$$

where c is a wave number.

Non-Stationarity across Longitudes:

Previous modified Matern:

$$f_L(c; \phi_L, \alpha_L, \nu_L) = \frac{\phi_L}{\left(\alpha_L^2 + 4\sin^2(c/N\pi)\right)^{\nu_L + 1/2}}, \quad c = 0, \cdots, N-1$$

$$\text{Here: } f_{L_m, l_n}(c) = \sum_{j=1}^{p} f_{L_m}^j(c) X^j(L_m, l_n)$$

In this work, land and ocean are included as covariates.

$$f_{L_m, l_n}(c) = f_{L_m}^1(c) b_{\text{ land }}(L_m, l_n) + f_{L_m}^2(c)\left\{1 - b_{\text{ land }}(L_m, l_n)\right\}$$

where

$$\left|f_{L_m}^j(c)\right|^2 = \frac{\phi_{L_m}^j}{\left\{\left(\alpha_{L_m}^j\right)^2 + 4\sin^2(c\pi/N)\right\}^{\nu_{L_m}^j + 1/2}}, \quad j = 1, 2$$

$$b_{\text{ land }}\left(L_m, l_n; g_{L_m}, \gamma_{L_m}\right) = \sum_{n'=1}^{N} \tilde{l}_m\left(l_n; g_{L_m}\right) w_m\left(l_n - l_{n'}; \gamma_{L_m}\right)$$

where $\tilde{l}_m\left(l_n; g_{L_m}\right)$ is equal to 1 for g_{L_m} more grid points wherever there is a land-ocean transition. And $w_m\left(l_n - l_{n'}; \gamma_{L_m}\right)$ is the Tukey taper function with range γ_{L_m},

$$w_m\left(l_n; \gamma_{L_m}\right) = \begin{cases} \frac{1}{2}\left[1 + \cos\left\{\frac{2\pi}{\gamma_{L_m}}\left(l_n - \frac{\gamma_{L_m}}{2}\right)\right\}\right], & 0 \leqslant \gamma_{L_m} < \frac{\gamma_{L_m}}{2} \\ 1, & \frac{\gamma_{L_m}}{2} \leqslant l_n < 1 - \frac{\gamma_{L_m}}{2} \\ \frac{1}{2}\left[1 + \cos\left\{\frac{2\pi}{\gamma_{L_m}}\left(l_n - 1 - \frac{\gamma_{L_m}}{2}\right)\right\}\right], & 1 - \frac{\gamma_{L_m}}{2} \leqslant l_n \leqslant 2\pi \end{cases}$$

Non-Stationarity across Latitudes

Previous Coherence:

$$\rho_{L_m,L_{m'}}(c) = \rho_{L_m-L_{m'}}(c) = \left[\frac{\xi}{\left\{1+4\sin^2(c\pi/N)\right\}^{\tau}}\right]^{|L_m-L_{m'}|} = \varphi(c)^{|L_m-L_{m'}|}$$

where $\varphi(c) = \frac{\xi}{\left\{1+4\sin^2(c\pi/N)\right\}^{\tau}}$

Here, authors propose a novel non-stationary model for the coherences, with latitudinally varying parameters

$$\varphi_{L_m}(c) = \frac{\xi_{L_m}}{\left\{1 + 4\sin^2(c\pi/N)\right\}^{\tau_{Lm}}}$$

Dataset

- Previous: CCSM3: 42×96 grids, 500 years, 5 realizations.
- Here: National Center for Atmosphere Research community climate system: 142×288 grids, 95 years, 6 realizations.

Model Fit

The model consists of three sets of parameters to be estimated

- Temporal Parameters, θ_{time}, consisting of all entries in $\mathbf{\Phi_1}$, $\mathbf{\Phi_2}$ and $\boldsymbol{S}$.
- Longitudinal Parameters, θ_{lon}, consisting of $\left(\phi^j_{L_m}, \alpha^j_{L_m}, \nu^j_{L_m}\right)$,$j = 1, 2$, and (g_{L_m}, γ_{L_m})
- Latitudinal Parameters, θ_{lat}, consisting of (ξ_{Lm}, τ_{L_m}).

Model Fit

- Step 1: Estimate the temporally AR parameters θ_{time}, assuming that the innovations $\boldsymbol{H}(t_k; r)$ are independent across latitude and longitude.

- Step 2: Consider θ_{time} fixed at their estimated values and estimate θ_{lon}, assuming that the innovations $\boldsymbol{H}(t_k; r)$ are independent across latitudes.

- Step 3: Consider θ_{time} and θ_{lon} fixed at their estimated values and estimate θ_{lat}.

Comments

- There are too many parameters to fit, more than 120,000 in step 1.
- Although it takes less than one day to fit 20 million data, it is much longer than that of the global model for climate ensembles.
- This improved model is much more complicated compared to the previous one.

Reference

- Castruccio, Stefano, and Joseph Guinness. "An evolutionary spectrum approach to incorporate largescale geographical descriptors on global processes." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 66, no. 2 (2017): 329-344.

- Jeong, Jaehong, Mikyoung Jun, and Marc G. Genton. "Spherical process models for global spatial statistics." *Statistical Science* 32, no. 4 (2017): 501-513.

- Castruccio, Stefano, and Michael L. Stein. "Global spacetime models for climate ensembles." *The Annals of Applied Statistics* 7, no. 3 (2013): 1593-1611.

Q & A

https://github.com/chxyself25/GLOBAL-SPACE-TIME-MODELS