Axially symmetric processes to model spatial data on large portions of the planet

Laboratorio de Modelación I

Laura Carrasco
Clemente Ferrer

Universidad Técnica Federico Santa Maria
Departament of Mathematics

August 10, 2022

Outline

Motivation

Dataset from *NOAA Physical Sciences Laboratory*.

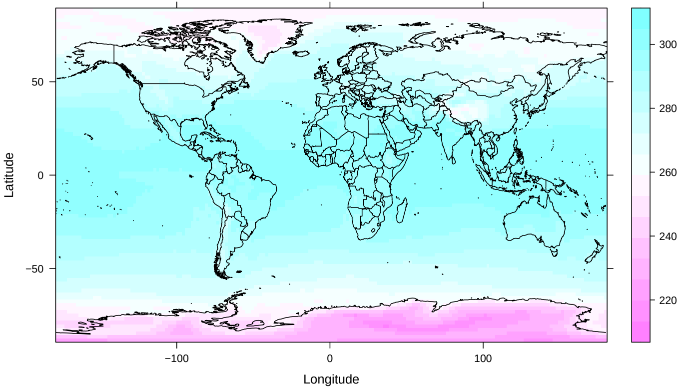

Figure: Air Temperature at 2m Above Ground Level (2021).

Motivation

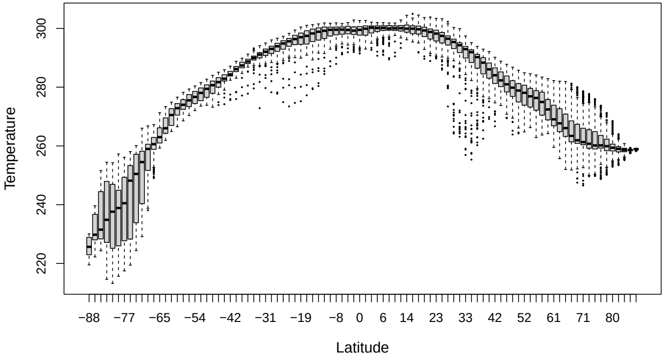

Figure: Air temperature by latitude band.

Random Fields

Random Fields. Let $\mathcal{D} \subset \mathbb{R}^d$ a spatial domain. A random field is a collection $\{Z(\mathbf{x}), \mathbf{x} \in \mathcal{D}\}$ of random variables.

In our framework we will work with the following assumptions

1. $\mathcal{D} \subset \mathbb{S}^2$ and $\mathbf{x} = (L, \ell)$, where L, ℓ denote latitude, in a range of $[0, \pi]$, and longitude is measured between $[0, 2\pi)$, respectively.
2. Second order randoms fields, i.e., $\mathrm{var}[Z(L, \ell)] < \infty$. Thus, mean and covariance function are well defined as follows

$$\mu(L, \ell) = E[Z(L, \ell)] \quad \text{and} \quad K(L_1, \ell_1, L_2, \ell_2) = \mathrm{cov}[Z(L_1, \ell_1), Z(L_2, \ell_2)].$$

Axially Symmetric Processes

Random field is referred to as an **axially symmetric process** if

1. $E[Z(L,\ell)] = m(L)$, for some function $m : [0,\pi] \to \mathbb{R}$.
2. $\operatorname{cov}\left[Z\left(L_1,\ell_1\right), Z\left(L_2,\ell_2\right)\right] = C\left(L_1, L_2, \ell_1 - \ell_2\right)$, for some function $C : \mathbb{S}^2 \to \mathbb{R}$.

But, **how do we estimate these functions?**

- m is straightforward given data.
- Covariance function must be positive definite, so we will estimate it using **Matérn Model**

$$C(\mathbf{h}) = \frac{2^{1-\nu}}{\Gamma(\nu)}(\alpha\|\mathbf{h}\|)^{\nu} K_{\nu}(\alpha\|\mathbf{h}\|), \quad \mathbf{h} \in \mathbb{S}^2,$$

where Γ, K_ν are gamma and modified Bessel functions of second kind.

Exponential Matérn Model and empirical covariance

In particular, if $\nu = 1/2$, then we obtain an **exponential Matérn model** given by

$$C(\mathbf{h}) = \exp(-\alpha\|\mathbf{h}\|).$$

Now, in order to estimate α, we will calculate residuals

$$\tilde{Z}(L,\ell) = \frac{Z(L,\ell) - \widehat{\mu}(L)}{\widehat{\sigma}(L)},$$

and fit a curve for empiral covariance

$$\widehat{C}(\mathbf{h}) = \frac{1}{N(\mathbf{h})} \sum_{(i,j):\|\mathbf{x}_i - \mathbf{x}_j\| \approx |\mathbf{h}|} Z(\mathbf{x}_i)Z(\mathbf{x}_j), \quad \mathbf{h} \in \mathbb{S}^2.$$

Mean and Standard Deviation Functions

Using **generalized additive model (GAM)** in R we predicted mean and standard deviation functions

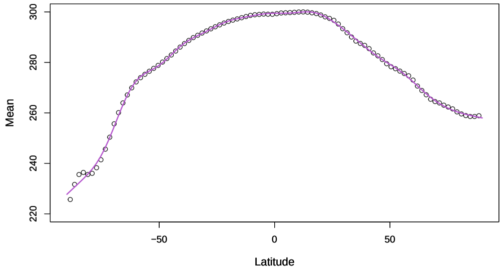

Figure: Mean function.

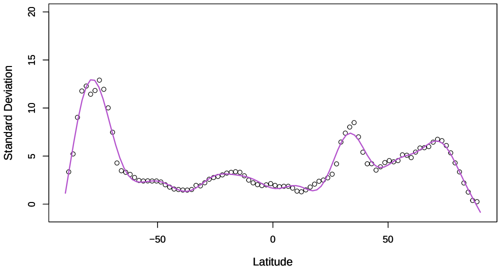

Figure: Standard deviation function.

Residues and normalization

Normalizing errors we obtain **residuals** of our data.

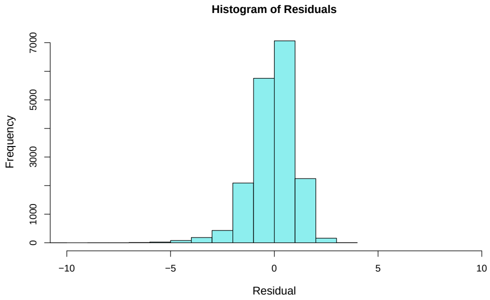

Figure: Histogram of residuals.

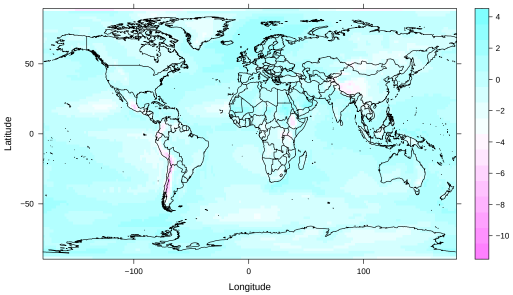

Figure: Residual worlwide map.

Residual Representation in Spherical Coordinates

Using spherical coordinates

$$\begin{cases} x = R\cos(L\omega)\cos(\ell\omega) \\ y = R\cos(L\omega)\sin(\ell\omega) \\ z = R\sin(L\omega) \end{cases}$$

where $R = 6378$ [km] and $\omega = \pi/180$.

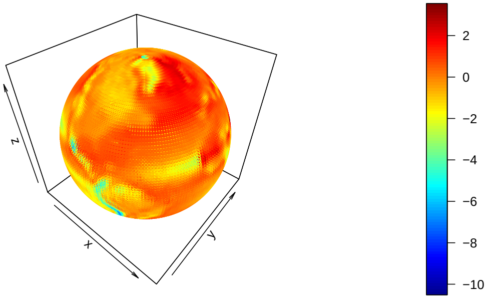

Figure: Residual projection over sphere.

Variogram and Empirical Covariance

Plotting empirical covariance and fitting a curve we obtain our model.

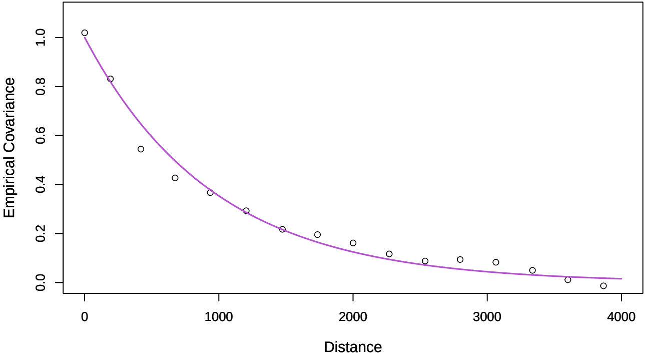

Figure: Empirical covariance.

Model applications: Kriging

- Method of interpolation based on Gaussian process governed by prior covariances.
- **Kriging** predicts the value of a function at a given point by computing a weighted average of the known values of the function in the neighborhood of the point.
- **Simple kriging**:

$$\mu(\mathbf{x}) = 0 \quad \text{and} \quad C(\mathbf{x}_1, \mathbf{x}_2) = \mathrm{cov}(Z(\mathbf{x}_1, \mathbf{x}_2)),$$

  for all $\mathbf{x} \in \mathbb{S}^2$.

Since residuals are normalized, $\mu \sim 0$, **simple kriging** can be applied because of we have a known covariance function.

Downscaling projections

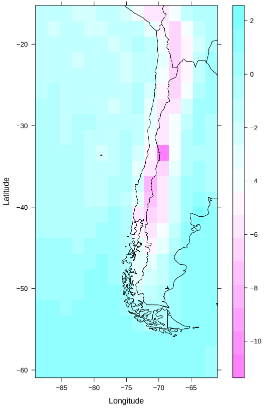

Figure: Original data.

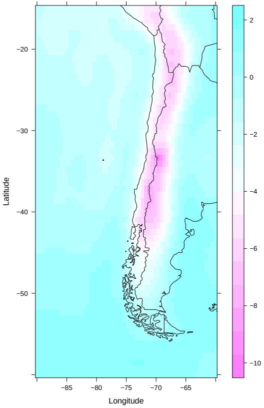

Figure: Kriging downscaling.

Conclusions and future work

Some remarkable facts about the project:

- Axially symmetric covariance functions technique was developed on $\mathbb{S}^2$.
- All the calculations were done using the euclidean distance on the sphere.
- Kriging method allowed us to build accurate temperature maps around Chile.

For future analysis, project files were uploaded to a public repository `MAT282`. Commendable future work would be to compare the predictions obtained with a different covariance model.

Bibliography

C. Gaetan and X. Guyon.
Spatial statistics and modeling.
Springer New York, 2010.

R. Olea.
Geostatistics for engineers and earth scientists.
Kluwer Academic, 1999.

E. Porcu, S. Castruccio, A. Alegría, and P. Crippa.
Axially symmetric models for global data: A journey between geostatistics and stochastic generators.
Environmetrics, 30(1), 2019.

M. Stein.
Spatial variation of total column ozone on a global scale.
Annals of Applied Statistics, 1:191–210, 2007.

Thank you.
Questions?