Anisotropy through spatial deformation in different space-time models

Marina S. Paez, Fidel E. C. Morales, Rodrigo Bulhões, Dimitris Politis, Jacek Leskow

November 3, 2020

Outline

1. Introduction
2. Student's-t process for space-time models with spatial deformation (Fidel E. C. Morales, Dimitris Politis, Jacek Leskow and Marina S. Paez)
   - Proposed model
   - Prior specifications and Inference procedures
   - Application
   - Conclusions
3. Spatial deformation in spatiotemporal models with multivariate responses (Rodrigo Bulhões, Marins S. Paez)
   - Motivation

Introduction

- Space-time models have been extensively explored in the previous 40 years to handle environmental phenomena, and they usually assume normality and isotropy;
- The urge to find better representations of reality lead researchers to explore more flexible approaches and relax these assumptions;
- In the first paper we propose a student's-t processes for cases when the data present heavier tails and/or discrepant values, and handle anisotropy through space deformation;
- In the second paper we propose a an anisotropic space-time model for multivariate data;

Student's-t process for space-time models with spatial deformation (Fidel E. C. Morales, Dimitris Politis, Jacek Leskow and Marina S. Paez)

- We propose the use of a Student's-t process, assuming spatial dependency through the specification of a spatial correlation function, and handling anisotropy through spatial deformation [Sampson and Guttorp, 1992];
- Evolution in time is done through Dinamic Models
- Working under the Bayesian approach to inference, we propose an efficient MCMC method to sample from the posterior distribution of the unknown parameters in the model;
- The proposed methodology is applied to a Hydrology dataset (for fixed time) and Pitch evaporation in Rio Grande do Sul (space-time).

Model

Consider a geographic region of interest denoted by $G \subset \mathbb{R}^p$, and suppose that a random process is observed at n fixed geographic points in G, at T distinct moments in time. Denote these locations by $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$.

Our observations are indexed by t, $t = 1, \ldots, T$, and denoted by $\boldsymbol{Y}_t = (Y(\boldsymbol{s}_1, t), \ldots, Y(\boldsymbol{s}_n, t))', t = 1, \ldots, T$.

When $p = 2$ we can define $\boldsymbol{s}_i = (x_i, y_i)'$,

We assume that the observations $\boldsymbol{Y} = (\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_T)$ can be explained by a sum of two independent components:

$$Y(\boldsymbol{s}_i, t) = \mu(\boldsymbol{s}_i, t) + W(\boldsymbol{s}_i, t), \; i = 1, \ldots, n, \; t = 1, \ldots, T,$$

with $\mu(\cdot, \cdot)$ representing the mean of the process and $W(\cdot, \cdot)$ representing smooth deviations from the mean.

In matrix notation, the model can be written as

$$\boldsymbol{Y} = \boldsymbol{\mu} + \boldsymbol{W}, \tag{1}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{W}$ are $n \times T$ matrices defined as $\boldsymbol{\mu} = (\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_T)$, with $\boldsymbol{\mu}_t = (\mu(\boldsymbol{s}_1, t), \ldots, \mu(\boldsymbol{s}_n, t))'$; and $\boldsymbol{W} = (\boldsymbol{W}_1, \ldots, \boldsymbol{W}_T)$, with $\boldsymbol{W}_t = (W(\boldsymbol{s}_1, t), \ldots, W(\boldsymbol{s}_n, t))'$.

Explanatory variables can be incorporated into the model through the mean process $\boldsymbol{\mu}$

A simple form:

$$\boldsymbol{\mu}(\boldsymbol{s}_i, t) = \boldsymbol{X}'_{i,t}\boldsymbol{\beta}. \tag{2}$$

We assume a stochastic variation in time through a state space model formulation.

More specifically, we have that $\boldsymbol{\mu}$ is a purely temporal component, such that $\mu(\boldsymbol{s}_1, t) = \ldots = \mu(\boldsymbol{s}_n, t)$, and we can write $\boldsymbol{\mu}_t = \mathbf{1}_n \mu_t$, $t = 1 \cdots, T$ as:

$$\begin{aligned} \text{Structural equation: } & \mu_t = \boldsymbol{F}_t' \boldsymbol{\beta}_t, \\ \text{System equation: } & \boldsymbol{\beta}_t = \boldsymbol{G}_t \boldsymbol{\beta}_{t-1} + \boldsymbol{w}_t, \boldsymbol{w}_t \sim N(\mathbf{0}, \boldsymbol{\Psi}), \end{aligned}$$

where $\boldsymbol{F}_t$ is a known matrix which can include covariates, and $\boldsymbol{G}_t$ is a known transition matrix.

We also assume that $\boldsymbol{W_t}'s$ are independent, for $t = 1, \ldots, T$, and all the temporal dependence is incorporated into the mean process.

The spatial dependence is introduced in the model through the specification of $\boldsymbol{W}_t' s$.

We assume a student's-t process with smooth variation in space for $\boldsymbol{W_t}$:

$$\boldsymbol{W_t} \sim t_\nu(\mathbf{0}, \boldsymbol{\Sigma}),$$

where

$$\boldsymbol{\Sigma} = \sigma_\xi^2 \boldsymbol{R}(\phi) + \sigma_\epsilon^2 \boldsymbol{I}_n,$$

- $\boldsymbol{R}(\phi)$ is a covariance matrix, such that $\boldsymbol{R}(\phi)[i,j] = \rho_\phi(\boldsymbol{s}_i, \boldsymbol{s}_j)$ for $i, j = 1, \ldots, n$, where ρ_ϕ is a valid correlation function.
- $\sigma_\epsilon^2 \boldsymbol{I}_n$ represents the nugget effect

Another way of representing $\boldsymbol{W}_t$ (Normal-Gamma representation)

Consider a latent vector $\boldsymbol{U} = (U_1, \ldots, U_T)$, such that $\boldsymbol{W}_t = U_t^{-1/2} \boldsymbol{Z}_t$, with
$\boldsymbol{Z}_t \overset{iid}{\sim} N_n(\mathbf{0}, \boldsymbol{\Sigma})$ and $U_t \overset{iid}{\sim} G(\nu/2, \nu/2)$.

That way, we can write:

$$\boldsymbol{Y}_t = \boldsymbol{\mu}_t + U_t^{-1/2} \boldsymbol{Z}_t, \qquad t = 1, \ldots, T. \tag{3}$$

Anisotropy

- To handle anisotropy, we follow the idea proposed by [Sampson and Guttorp, 1992]:
- Considers a function $\boldsymbol{d}(\cdot)$ to map the original geographic coordinates from space G to the new space $D \subset \mathbb{R}^q$, $q \geq p$ (we assume $q = 2$), where the hypothesis of isotropy hold.
- Particularly, we work with the exponential correlation function $\rho_\phi(\boldsymbol{s}_i, \boldsymbol{s}_j) = e^{-\phi|\boldsymbol{d}_i - \boldsymbol{d}_j|}$.

Coordinates in D-space

We assume for $\boldsymbol{d}(\cdot)$ the prior specification proposed by [Schmidt and O'Hagan, 2003], given by:

$$\boldsymbol{d}(\cdot) \sim GP(\boldsymbol{g}(\cdot), \boldsymbol{\sigma}_{\boldsymbol{d}}^{\boldsymbol{2}} \rho_{\psi}(\cdot)), \tag{4}$$

- Particularly, we assume the identity function $\boldsymbol{g}(\boldsymbol{s}) = \boldsymbol{s}$.
- We assume $\boldsymbol{\sigma}_{\boldsymbol{d}}^{\boldsymbol{2}}$ to be a 2×2 diagonal matrix, which controls the variance of the deviation in each coordiate from the original space.
- $\rho_{\psi}(\cdot)$ is defined as Gaussian correlation function, such that $\rho_{\psi}(x) = \exp\{-\psi x^2\}$.
- We consider ψ to be known.

Anisotropy

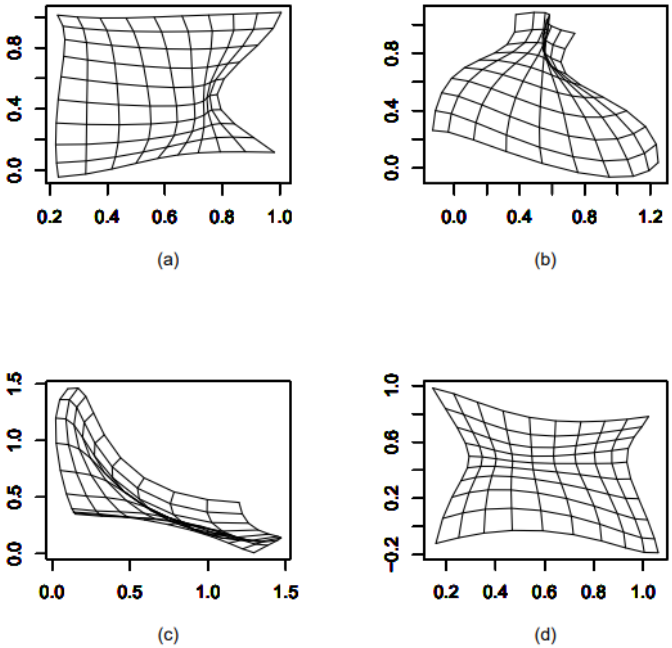


Figure 1: Deformations of the regular grid $[0,1] \times [0,1]$ generated with parameters $\phi = 3.7$, a) $\sigma^2_{d_{1,1}} = \sigma^2_{d_{2,2}} = 0.01$; b) $\sigma^2_{d_{1,1}} = \sigma^2_{d_{2,2}} = 0.03$; c) $\sigma^2_{d_{1,1}} = \sigma^2_{d_{2,2}} = 0.05$; d) $\sigma^2_{d_{1,1}} = 0.01$, $\sigma^2_{d_{2,2}} = 0.05$.

Prior distribution for degrees of freedom parameter ν

The prior distribution for parameter ν is the same proposed by [Cabral et al., 2012], that is:

$$\nu|\lambda \sim G(1, \lambda)$$
$$\lambda \sim U(a_\lambda, b_\lambda)$$

Prior distribution for the parameters σ_ξ^2, σ_ϵ^2 and ϕ

- prior distribution proposed by [Schmidt and O'Hagan, 2003] and [Castro Morales et al., 2013] for ϕ, that is:

  $\phi \sim G(a_\phi \eta, \eta)$, where $a_\phi = -2\log(0.05)/max(\mid \mathbf{s}_i - \mathbf{s}_j \mid)$.

- The prior distributions for the variances σ_ξ^2, σ_ϵ^2 are based on an alternative representation proposed by [Yan et al., 2007] for matrix $\mathbf{\Sigma}$.

  They propose writing

$$\mathbf{\Sigma} = \sigma^2[(1-\kappa)\boldsymbol{R}(\phi) + \kappa \boldsymbol{I}_n],$$

  such that $\sigma^2 = \sigma_\xi^2 + \sigma_\epsilon^2$ and $\kappa = \frac{\sigma_\epsilon^2}{\sigma^2}$. Thus, κ represents the fraction of the total variability of $\boldsymbol{Y}$ that corresponds to the measurement error.

The prior distributions of σ^2 and κ proposed by [Yan et al., 2007] are

$$\pi(\kappa) \propto \frac{\kappa^{a_\xi - 1}(1-\kappa)^{a_\epsilon - 1}}{[b_\xi \kappa + b_\epsilon(1-\kappa)]^{a_\xi + a_\epsilon}}, \ \kappa \in (0,1), \tag{5}$$

$$\sigma^2 \mid \kappa \sim GI\left(a_\epsilon + a_\xi, \frac{b_\xi}{1-\kappa} + \frac{b_\epsilon}{\kappa}\right). \tag{6}$$

These prior distributions are induced by the prior distributions for σ_ξ^2 and σ_ϵ^2, given by:

$\sigma_\xi^2 \sim IG(a_\xi, b_\xi)$ and $\sigma_\epsilon^2 \sim IG(a_\epsilon, b_\epsilon)$,

where $IG(a_{IG}, b_{IG})$ denotes the inverse Gamma distribution with mean $b_{IG}/(a_{IG} - 1)$.

regression coefficent $\boldsymbol{\beta}$

We assign a normal conjugate prior, given by

$$\boldsymbol{\beta} \sim N(\boldsymbol{M}_0, \boldsymbol{C}_0),$$

with known values of $\boldsymbol{M}_0$ and $\boldsymbol{C}_0$.

Define $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \nu, \lambda, \sigma^2, \kappa, \phi, \boldsymbol{d}^G, \boldsymbol{\sigma}_{\boldsymbol{d}}^{\boldsymbol{2}}\}$ as the complete set of unknown model parameters.

- Prior distribution:

$$\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta})\pi(\nu \mid \lambda)\pi(\lambda)\pi(\sigma^2)\pi(\kappa)\pi(\phi)\pi(\boldsymbol{d}^G \mid \boldsymbol{\sigma}_{\boldsymbol{d}}^{\boldsymbol{2}})\pi(\boldsymbol{\sigma}_{\boldsymbol{d}}^{\boldsymbol{2}})$$

- Likelihood:

$$L(\boldsymbol{\theta} \mid \boldsymbol{Y}) = \prod_{t=1}^{k} \frac{\Gamma(\frac{p+\nu}{2})}{\Gamma(\frac{\nu}{2})\pi^{p/2}} \nu^{-p/2} \mid \boldsymbol{\Sigma} \mid^{-1/2} \left(1 + \frac{d(\boldsymbol{Y}, \boldsymbol{\theta})}{\nu}\right)^{-\frac{p+\nu}{2}},$$

where $d(\boldsymbol{Y}, \boldsymbol{\theta}) = (\boldsymbol{Y} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{\mu})$.

The posterior distribution do not have a closed form and MCMC methods were proposed. Extra steps can be added to the algorithm to sample from predictive distributions.

Water heights in Saratoga Valley

- We apply the proposed model to the observation of water heights (in meters above sea level) at 93 fixed locations in a aquifer of 1300 square kilometers near Saratoga Valley, Wyoming ([Lenfest Jr, 1986]).

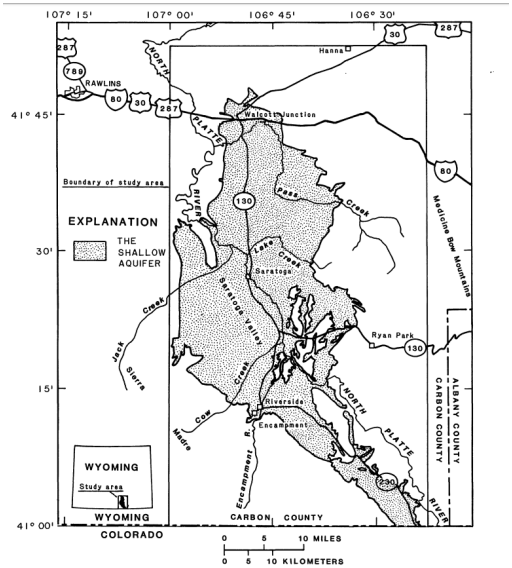


Figure 2: Location of the shallow aquifer within the study area.

Application

- One challengeful aspect of these observations is that the residuals obtained after explaining the data by a linear combination of the geographic coordinates present atypical values.
- Also, the semivariogram of these residuals suggest that the process is anisotropic.
- Besides the proposed model, which we will refer to as Model A, we also estimated for comparison an isotropic version of this model (Model B), and gaussian versions either considering spatial deformation (model C) or isotropy (Model D).
- We considered the geographic coordinates as explanatory variables.

The prior distribution of the model parameters were specified as:

- $\boldsymbol{\beta} \sim \mathrm{N}(0, 1000\boldsymbol{I})$,
- $\boldsymbol{\Psi} \sim \mathrm{WI}_{n_0}^{-1}(\boldsymbol{S}_0)$, where $n_0 = 2$ and $\boldsymbol{S}_0 = diag(1e^{-3}, 1e^{-3}, 1e^{-3})$,
- $\lambda \sim U(0.01, 1)$,
- $\phi \sim \mathrm{G}(2 \times 0.1, 0.1)$,
- $\sigma^2_{d_j} \sim \mathrm{GI}(2.01, 0.101)$, $j = 1, 2$;
- The priors for κ and $\sigma^2 \mid \kappa$ are as in (6) with $a_e = a_z = \frac{209}{100}$, $b_e = b_z = 3(a_e - 1)$.

Samples from the posterior distributions were obtained under each model through MCMC. 100,000 iterations of the MCMC algorithms were obtained, with a burn-in of 50,000.

95% credibility intervals, posterior means and medians obtained through MCMC samples for the parameters β_0, β_1, β_2 σ^2, κ, ϕ and ν under models A, B, C and D.

	Model A				Model B			
Par	Mean	50%	2.5%	97.5%	Mean	50%	2.5%	97.5%
β_0	2237.23	2237.35	2191.16	2283.86	2241.43	2241.29	2192.96	2290.07
β_1	0.93	0.94	-1.13	2.92	0.10	0.11	-1.82	2.11
β_2	-3.43	-3.42	-4.58	-2.26	-3.19	-3.19	-4.35	-2.03
σ^2	299.65	159.12	23.69	1402.04	337.23	201.73	33.74	1373.88
κ	0.01	0.01	0.00	0.40	0.01	0.01	0.00	0.03
ϕ	0.03	0.03	0.01	0.07	0.04	0.04	0.01	0.09
ν	3.84	2.92	2.03	11.57	4.52	3.12	2.04	16.22
DIC	-349.58				-312.16			
	Model C				Model D			
Par	Mean	50%	2.5%	97.5%	Mean	50%	2.5%	97.5%
β_0	2240.55	2240.45	2195.73	2285.91	2241.47	2241.46	2196.59	2286.46
β_1	0.79	0.78	-0.89	2.48	0.03	0.02	-1.77	1.85
β_2	-3.45	-3.44	-4.52	-2.34	-3.19	-3.19	-4.18	-2.20
σ^2	798.73	715.63	413.27	1686.47	815.99	721.08	419.35	1768.29
κ	0.02	0.01	0.00	0.05	0.01	0.01	0.00	0.04
ϕ	0.05	0.04	0.02	0.08	0.06	0.06	0.02	0.10
DIC=	-307.05				-278.63			

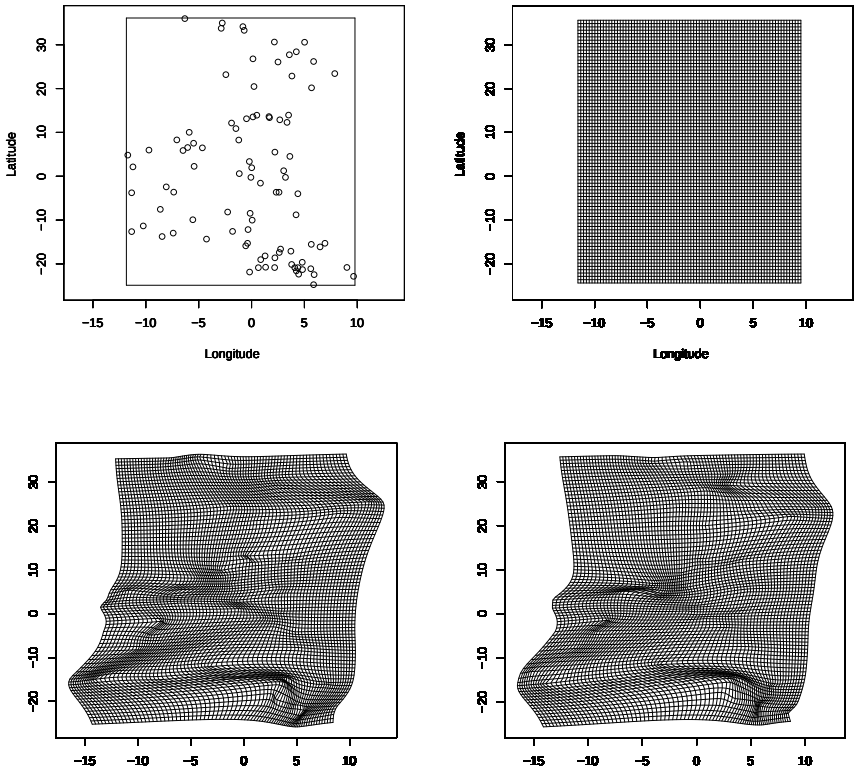


Figure 3: a) 93 fixed locations in the aquifer where observations were made; b) constructed 40×40 grid in region G for interpolation. c) Posterior mean of the spatial deformation under model A; d) Posterior mean of the spatial deformation under model C.

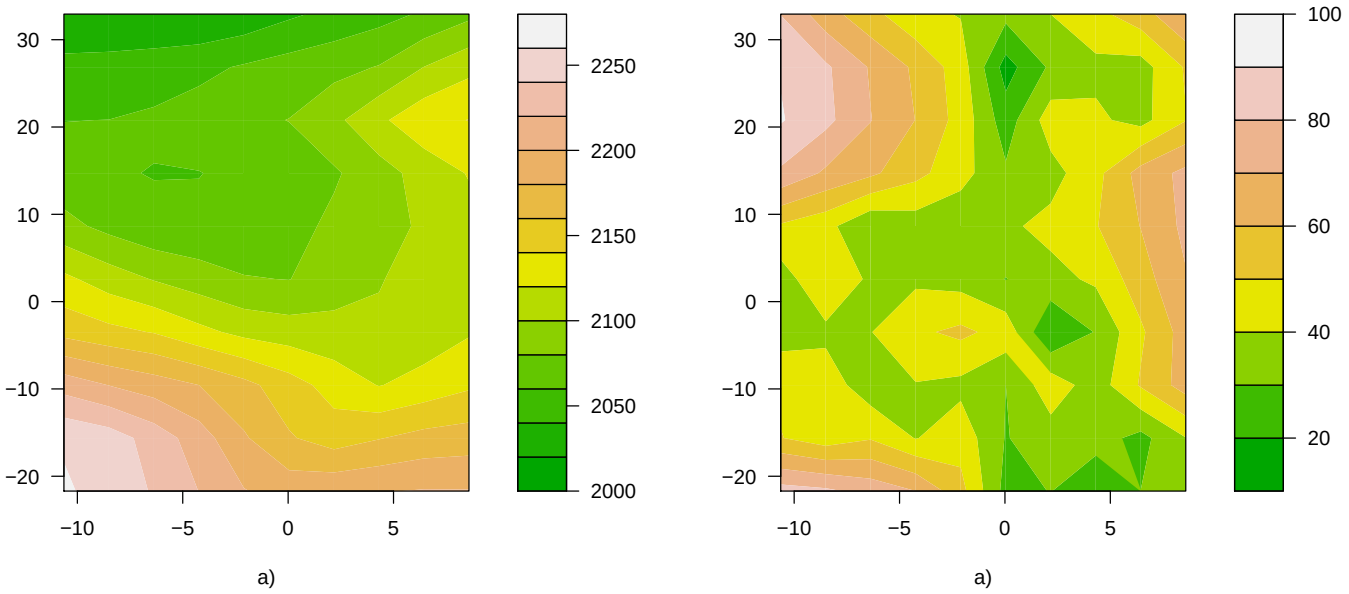
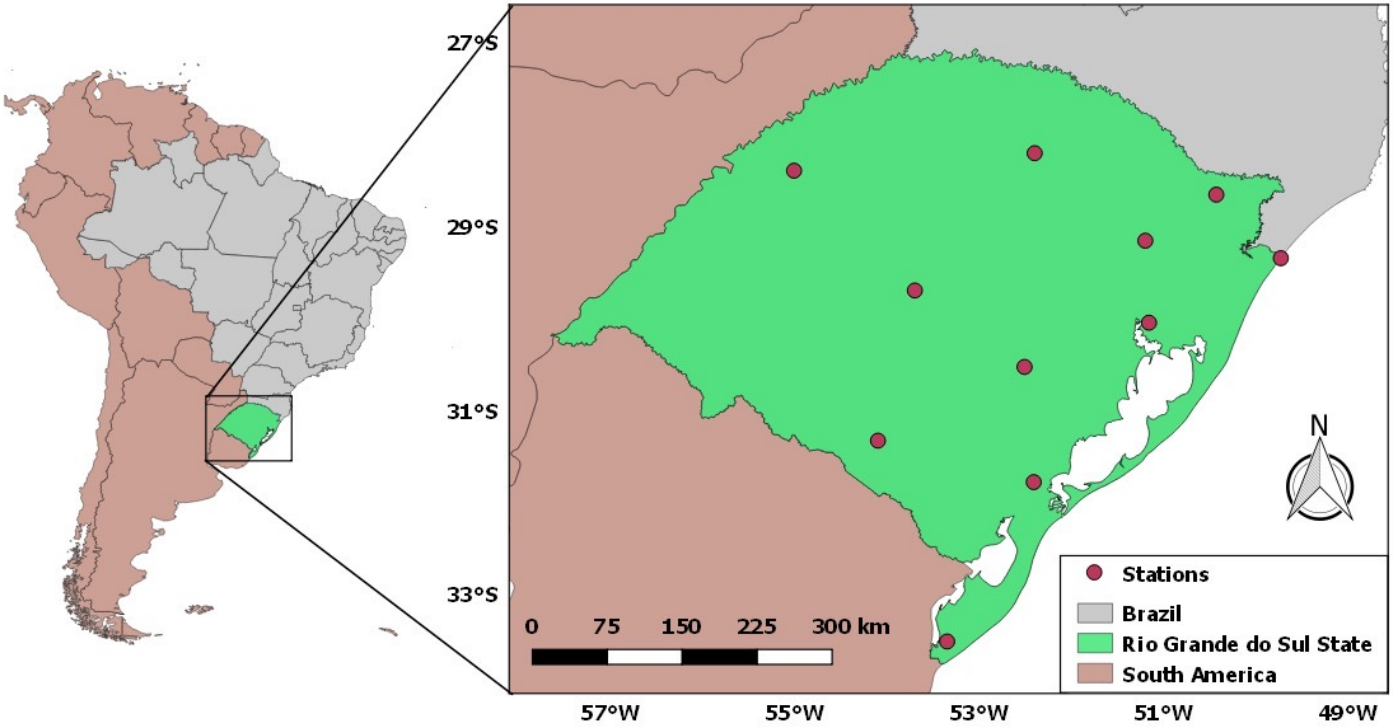


Figure 4: a) average a posteriori forecast of water height under sea level using Model A. b) Amplitude of the credibility range of 95% of water height above sea level using Model A.

Pitch evaporation in Rio Grande do Sul

- Pitch evaporation was measured daily at 11 stations in the state of Rio Grande do Sul (years 2017 and 2018, for a total T =730 observations per monitoring station).
- Exploratory analysis of the pitch evaporation time series show some atypical values
- The heterogenic topography of the region under study, as well as the climatic variations produced by the geographical location (south latitude) and the cold fronts coming from the south pole (Reboita et al., 2010) are an indication that the hypothesis of isotropy would not be reasonable for the pitch evaporation process.


27°S
29°S
31°S
33°S
0 75 150 225 300 km
57°W
55°W
53°W
51°W
49°W
N
Stations
Brazil
Rio Grande do Sul State
South America

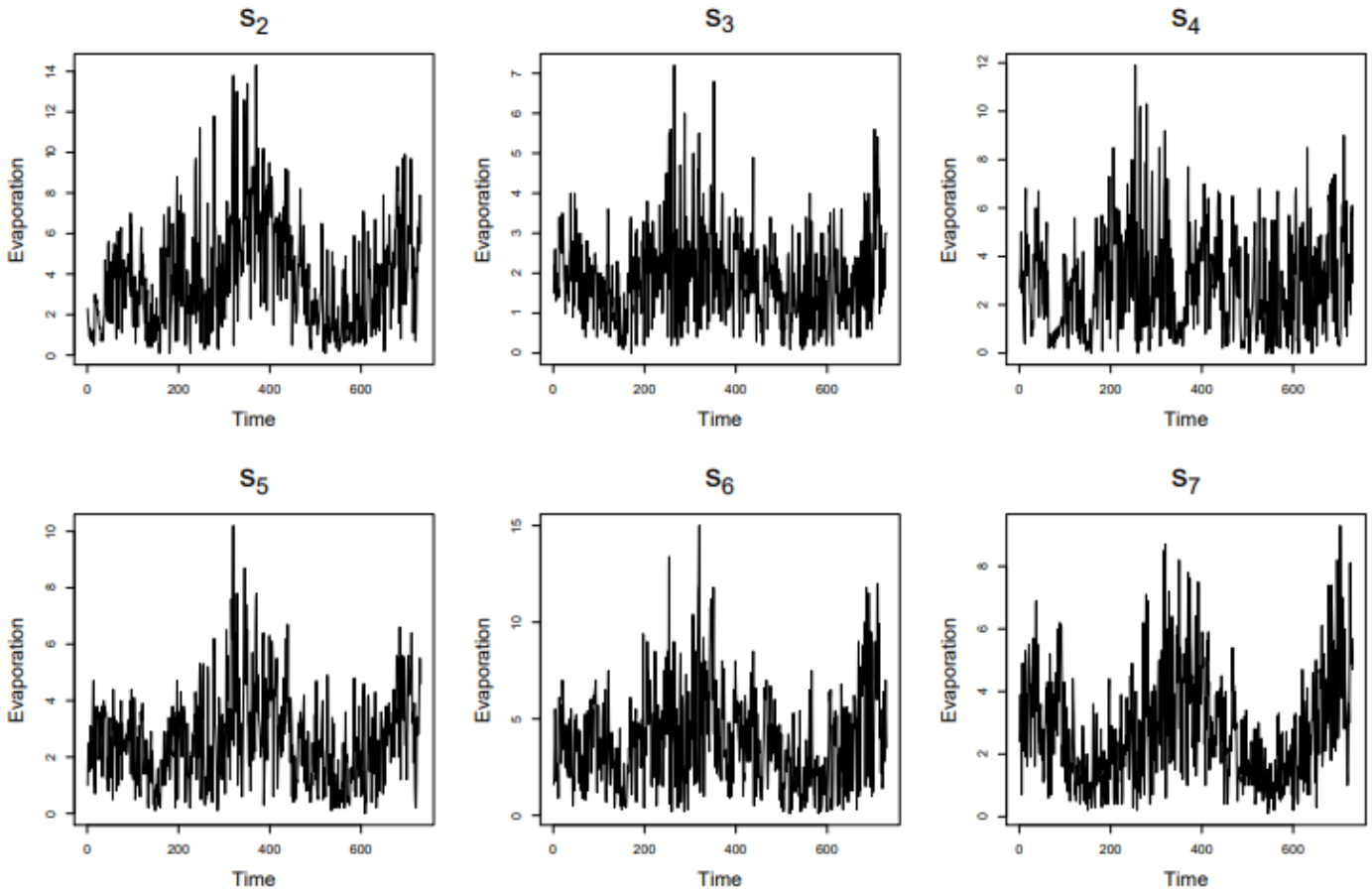


Figure 6: Measurements of Evaporation through time in 6 Monitoring Stations

Pitch evaporation in Rio Grande do Sul

We consider the following covariates to explain the mean of the process: **Latitude** (denoted by x_i); **Longitude** (denoted by y_i); **Altitude** (denoted by z_i), the **iterations** between x_i, y_i, and z_i; and an **annual and semiannual seasonal feat**. That is, the state space model is defined with the covariate vector

$\boldsymbol{F_{t}}_i = (1\ x_i\ y_i\ z_i\ x_iy_i\ x_iz_i\ y_iz_i\ x_iy_iz_i\ 1\ 0\ 1\ 0)$, for $s_i,\ i = 1, \ldots, 11$. And the transition matrix $\boldsymbol{G}_t = \begin{pmatrix} \boldsymbol{I}_8 & \boldsymbol{0}_{8\times 4} \\ \boldsymbol{0}_{4\times 8} & \boldsymbol{G}_3 \end{pmatrix}$, where
$\boldsymbol{G}_1 = \begin{pmatrix} \cos(\omega) & \sin(\omega) \\ -\sin(\omega) & \cos(\omega) \end{pmatrix}$, $\boldsymbol{G}_2 = \begin{pmatrix} \cos(2\omega) & \sin(2\omega) \\ -\sin(2\omega) & \cos(2\omega) \end{pmatrix}$,
$\boldsymbol{G}_3 = \mathrm{diag}(\boldsymbol{G}_1, \boldsymbol{G}_2)$, $\omega = 2\pi/365$.

Pitch evaporation in Rio Grande do Sul

We compare the following models:

- Model A: $\boldsymbol{W}_t \sim t_n(\mathbf{0}, \boldsymbol{\Sigma})$, $\Sigma_{ij} = \sigma^2[(1-\kappa)\phi\frac{\sin(|\boldsymbol{d}_i - \boldsymbol{d}_j|/\phi)}{|\boldsymbol{d}_i - \boldsymbol{d}_j|} + \kappa]$.
- Model B: $\boldsymbol{W}_t \sim t_n(\mathbf{0}, \boldsymbol{\Sigma})$, $\Sigma_{ij} = \sigma^2[(1-\kappa)\phi\frac{\sin(|\boldsymbol{s}_i - \boldsymbol{s}_j|/\phi)}{|\boldsymbol{s}_i - \boldsymbol{s}_j|} + \kappa]$,
- Model C: $\boldsymbol{W}_t \sim N_n(\mathbf{0}, \boldsymbol{\Sigma})$, $\Sigma_{ij} = \sigma^2[(1-\kappa)\phi\frac{\sin(|\boldsymbol{d}_i - \boldsymbol{d}_j|/\phi)}{|\boldsymbol{d}_i - \boldsymbol{d}_j|} + \kappa]$,
- Model D: $\boldsymbol{W}_t \sim N_n(\mathbf{0}, \boldsymbol{\Sigma})$, $\Sigma_{ij} = \sigma^2[(1-\kappa)\phi\frac{\sin(|\boldsymbol{s}_i - \boldsymbol{s}_j|/\phi)}{|\boldsymbol{s}_i - \boldsymbol{s}_j|} + \kappa]$.

We used the following prior distributions for the model parameters:

- $\boldsymbol{\Psi} \sim WI_{n_0}^{-1}(\boldsymbol{S}_0)$, em que $n_0 = 2$ e $\boldsymbol{S}_0 = I_{12}$,
- $\lambda \sim U(1/50, 1/3)$,
- $\phi \sim G(0.19, 1)$,
- $\pi(\kappa) \propto \frac{\kappa^{a_z-1}(1-\kappa)^{a_e-1}}{[b_z\kappa+b_e(1-\kappa)]^{a_z+a_e}}$,
- $\sigma^2 \mid \kappa \sim GI(a_e + a_z, \frac{b_z}{1-\kappa} + \frac{b_e}{\kappa})$, $a_e = a_z = \frac{209}{100}$, $b_e = b_z = 3(a_e - 1)$,
- $\boldsymbol{\beta}_0 \sim N(\mathbf{0}, 1000\boldsymbol{I})$,

We sampled from the posterior distribution of the parameters through MCMC, considering a sample of $10,000$ iterations obtained after a burn-in of $100,000$ iterations.

Table 1: Posterior mode, mean, median, and 95 % credibility interval for the parameters σ^2, ϕ, κ, and ν of the models adjusted for evaporation data in the states of Rio Grande do Sul in Brazil.

	Model A					Model B				
Par	Mean	Median	2.5%	97.5%	SD	Mean	Median	2.5%	97.5%	SD
σ^2	0.65	0.65	0.59	0.72	0.03	0.75	0.74	0.64	0.91	0.07
ϕ	0.62	0.61	0.57	0.67	0.03	6.07	5.95	2.52	10.61	2.08
κ	0.43	0.43	0.38	0.48	0.02	0.67	0.67	0.54	0.78	0.07
ν	4.25	4.24	3.58	4.93	0.35	4.04	4.04	3.47	4.69	0.31
DIC	1780					2075				
	Model C					Model D				
Par	Mean	Median	2.5%	97.5%	SD	Mean	Median	2.5%	97.5%	SD
σ^2	1.23	1.23	1.15	1.31	0.04	1.34	1.33	1.03	1.60	0.14
ϕ	0.58	0.58	0.55	0.60	0.01	6.48	6.51	1.09	10.78	2.14
κ	0.28	0.28	0.23	0.34	0.03	0.71	0.71	0.58	0.87	0.07
DIC	4201					5380				

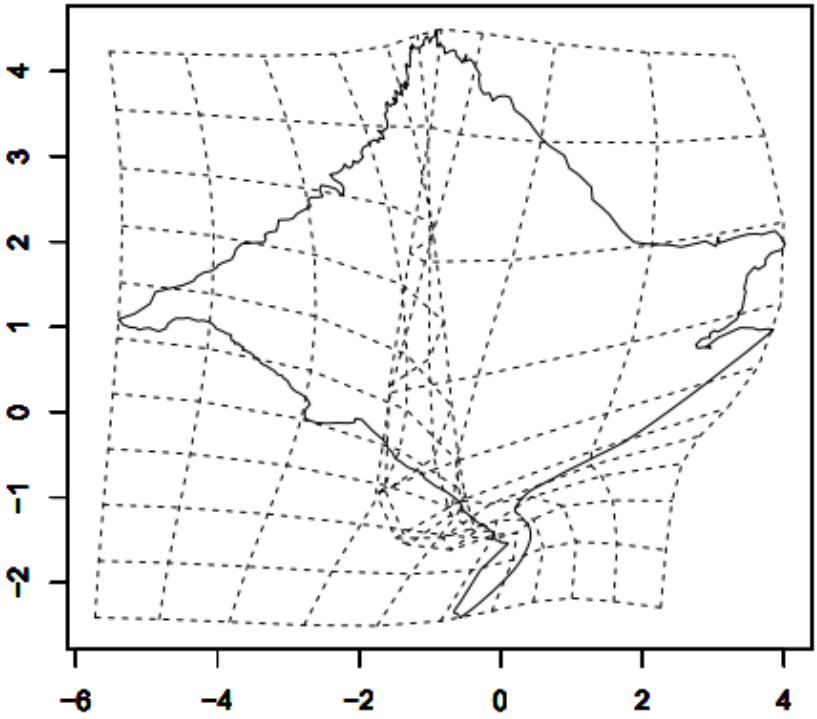


Figure 7: Deformed map of the state of Rio Grande do Sul.

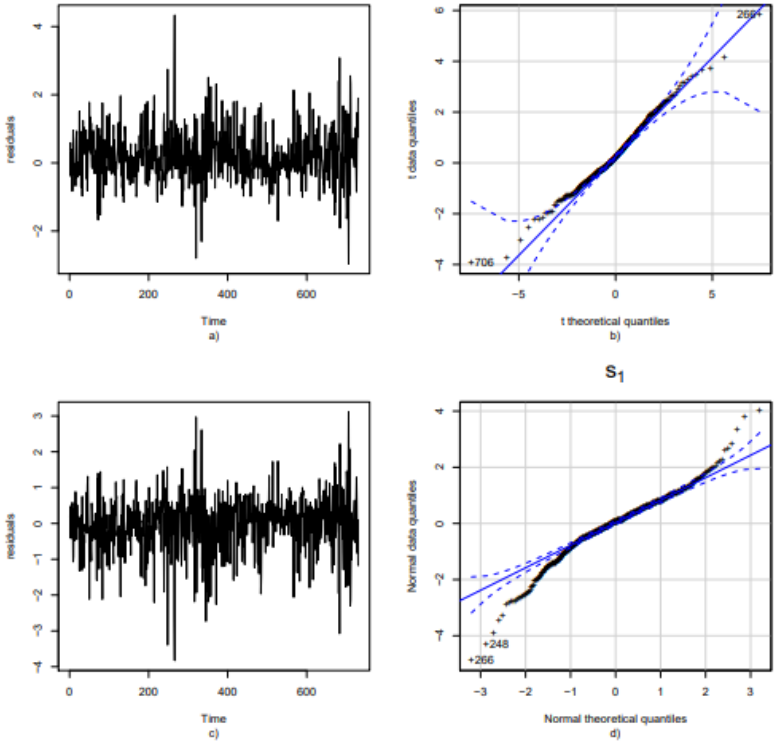


Figure 8: a) Residuais estimados pelo Modelo A para a estação s_1. b) Q-q plot of the t distribution with simulated confidence envelopes (95%) for the estimated residual with the Model A. c) Residuais estimados pelo Modelo C para a estação s_1. d) Q-q plot of the normal distribution with simulated confidence envelopes (95%) for the estimated residual with the Model C.

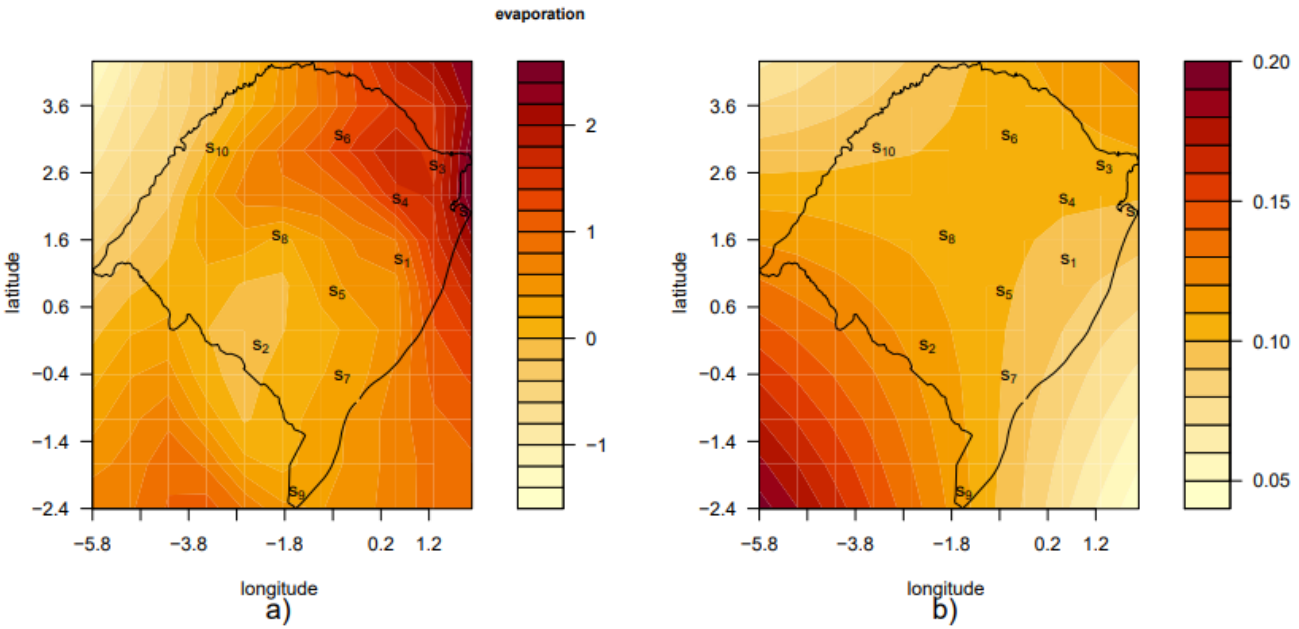


Figure 9: a) Superficie da média e b) superficie da amplitude do intervalo de credibilidade do 95% para evaporização de piche no Rio Grande do sul, para t=160.

Conclusions

- The proposed methodology performed better through the DIC criteria. Thus, this approach was shown to be a good alternative to model processes that generate observations with heavy tailed distributions or atypical data.

- As a future line of work, we would like to explore other approaches for heavy tailed distributions with atypical observations, such as the family of elliptical distributions proposed by [De Bastiani et al., 2015].

Spatial deformation in spatiotemporal models with multivariate responses (Rodrigo Bulhões, Marins S. Paez)

- In several fields of science, observations of different time series are observed at different locations in space.
- As an example, air quality indexes have been gauged daily at various stations located in the city of Rio de Janeiro, Brazil, to monitor the atmospheric pollution.
- Paez et al (2008) presented a class of linear models to treat measurements made in discrete time and continuous space, under the assumptions of stationarity and isotropy.
- Here we work under this same class of models but relaxing the hypothesis of isotropy through space deformation.

Univariate case

- Consider a set of discrete periods of time $t = 1, \ldots, T$, where for each t a random process $Y_t(\cdot)$ is observed in a set of locations in space $S = \{s_1, \ldots, s_N\}$

- Let $X_t(s)$ be a p-dimensional vector of covariates observed in time t and location s.

- Suppose $Y_t(\cdot)$ can be modeled through a regression equation, where the effects of the covariates vary smoothly along time and space as specified below:

$$
\begin{aligned}
Y_t(s) &= X_t'(s)\theta_{1,t}(s) + \nu_{1t}(s), \qquad \nu_{1t}(\cdot) \overset{iid}{\sim} N(0, V_1) \\
\theta_{1,t}(s) &= \theta_{2,t} + \nu_{2t}(s), \qquad \nu_{2t}(\cdot) \overset{iid}{\sim} f_\phi(\nu_{2t}), \\
\theta_{2,t} &= \theta_{2,t-1} + w_t, \qquad w_t \overset{iid}{\sim} N(0, W),
\end{aligned}
$$

for $t = 1, \ldots, T$ and $s = s_1, \ldots, s_N$

- Suppose that $f_\phi(\nu_{2t})$ defines a spatial correlation structure for these errors, and consequently for $\theta_{1,t}(\cdot)$.
- Supposing no spatial structure in V_1, the spatial dependence is obtained only through the regression parameters $\theta_{1,t}$.
- Many definitions for $f_\phi(\nu_{2t})$ are possible., for example: $f_\phi(\nu_{2t}) = N(0, V\rho(\phi, \cdot))$.
- Matrix V defines the covariance between coefficients which correspond to different covariates. These coefficients share the same spatial correlation defined through $\rho(\phi, \cdot)$.

- The correlation function ρ can be defined by any function that is able to capture the spatial correlation structure given in the data.
- Working under the hypothesis of isotropy, this function typically depends on d_{ij} , the distance between locations s_i and s_j, and parameters ϕ.
- Now we propose relaxing the hipothesis of isotropy and work under the hypothesis of isotropy in the new transformed space D.

Another way of describing this model is through a vector notation, given by

$$\begin{aligned} Y_t &= F_{1t}\theta_{1,t} + \nu_{1t}, & \nu_{1t} &\overset{iid}{\sim} N(0, V_1) \\ \theta_{1,t} &= F_{2t}\theta_{2,t} + \nu_{2t}, & \nu_{2t} &\overset{iid}{\sim} f_\phi(\nu_{2t}), \\ \theta_{2,t} &= G_t\theta_{2,t-1} + w_t, & w_t &\overset{iid}{\sim} N(0, W) \end{aligned}$$

for $t = 1, \ldots, T$.

- $\theta_{1,t}$ is a Np-dimensional vector, $\theta_{2,t}$ is a p-dimensional vector, the matrix F_{1t} is $N \times Np$, F_{2t} is $Np \times p$ and G_t is $p \times p$.
- The matrices F_{1t} , F_{2t} and G_t are assumed known, with F_{1t} and F_{2t} incorporating covariate values.
- $F_{1t} = diag(X_t(s_1), \ldots, X_t(s_N))$, $F_{2t} = 1_N I_p$ and $G_t = I_p$
- the spatial correlation presented in the response variables is given only through the regression parameters $\theta_{1,t}$, specifying $V_1 = \sigma^2_{\nu 1t} I_N$

Coordinates in D-space

We assume for $\boldsymbol{d}(\cdot)$ the prior specification proposed by [Schmidt and O'Hagan, 2003], given by:

$$\boldsymbol{d}(\cdot) \sim GP(\boldsymbol{g}(\cdot), \boldsymbol{\sigma}_{\boldsymbol{d}}^{\boldsymbol{2}} \rho_{\psi}(\cdot)), \tag{7}$$

- Particularly, we assume the identity function $\boldsymbol{g}(\boldsymbol{s}) = \boldsymbol{s}$.
- We assume $\boldsymbol{\sigma}_{\boldsymbol{d}}^{\boldsymbol{2}}$ to be a 2×2 diagonal matrix, which controls the variance of the deviation in each coordiate from the original space.
- $\rho_{\psi}(\cdot)$ is defined as Gaussian correlation function, such that $\rho_{\psi}(x) = \exp\{-\psi x^2\}$.
- We consider ψ to be known.

The model is completed with:

- independent inverted Wishart (Dawid, 1981) prior distributions for the covariance matrix W;
- inverted gamma prior distribution for $\sigma^2_{\nu 1t}$;
- normal prior distribution for $\theta_{2,0}$;
- prior distribution proposed by [Schmidt and O'Hagan, 2003] for ϕ: $\phi \sim G(a_\phi \eta, \eta)$, where $a_\phi = -2\log(0.05)/max(\mid \mathbf{s}_i - \mathbf{s}_j \mid)$.

Motivation of the multivariate case

- We wish to model measurements of: $PM_{2.5}$ (particulate matter with less than $2.5mg/m^3$); PM_{10} (particulate matter with less than $10mg/m^3$); and TSP (total amount of suspended particles).
- The observations were made in twenty-four semi automatic monitoring stations located in the Metropolitan region of Rio de Janeiro, Brazil, at every six days, from the year 2000 to 2015.
- The air quality indexes were observed simultaneously only in three monitoring stations. In the other monitoring stations only either one or two indexes of pollutants were measured.

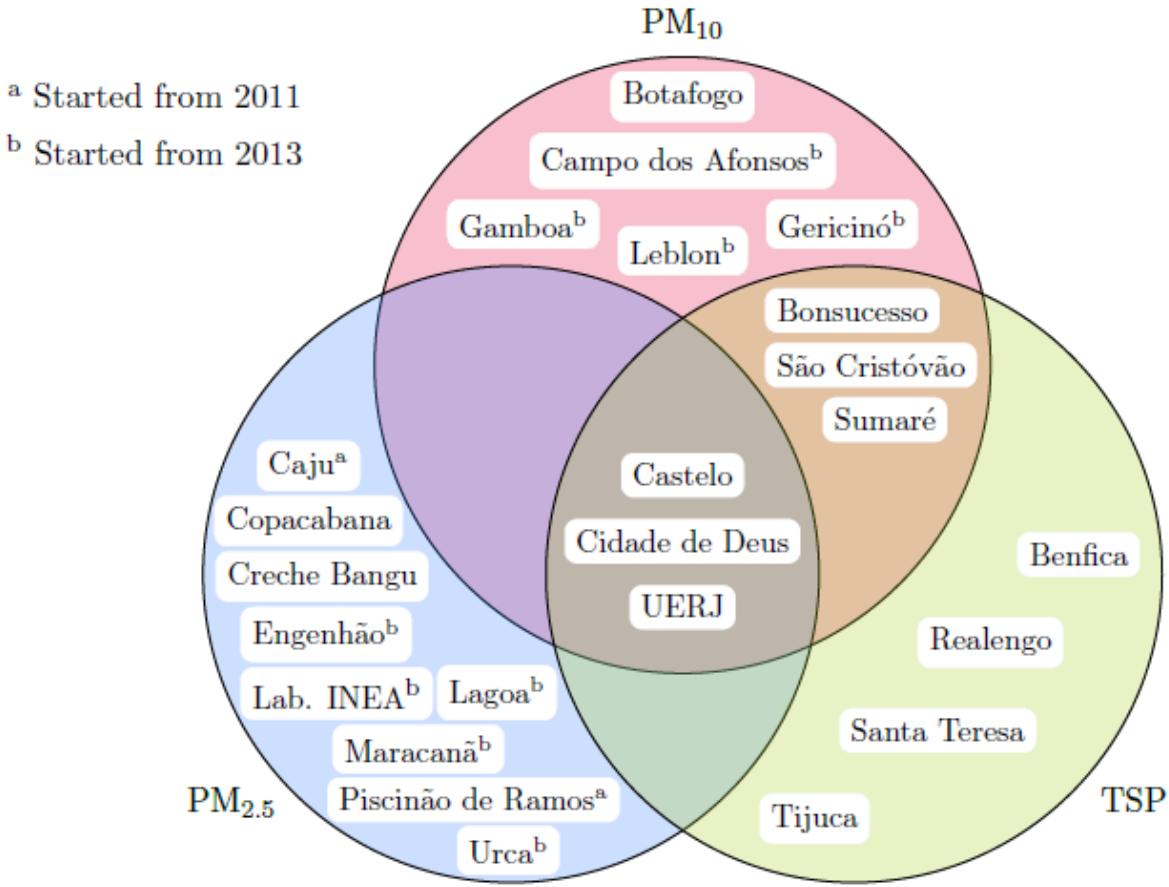


Figure 1.1: Monitoring stations associated to $PM_{2.5}$, PM_{10} and TSP.

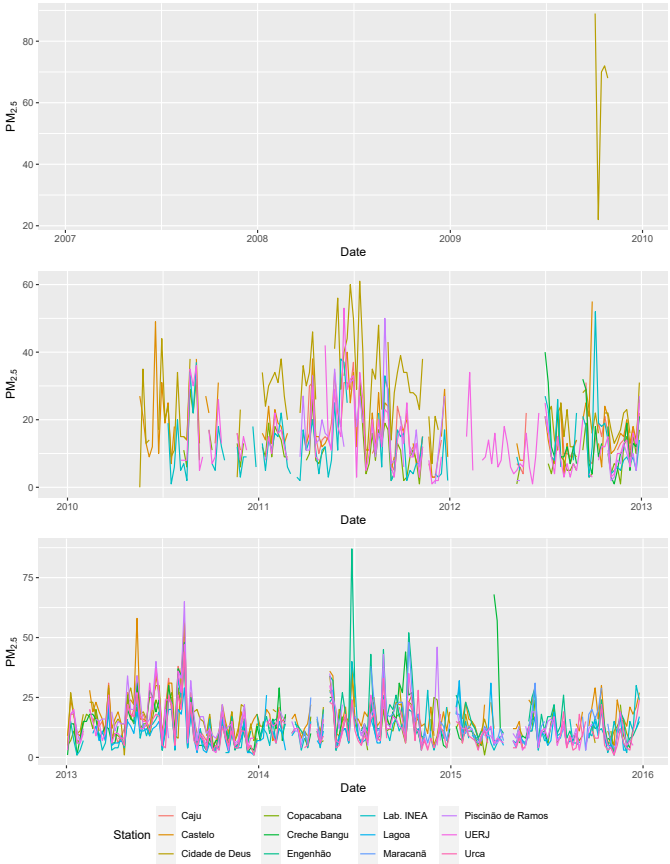


Figure 10: Time series of $PM_{2.5}$ per monitoring station.

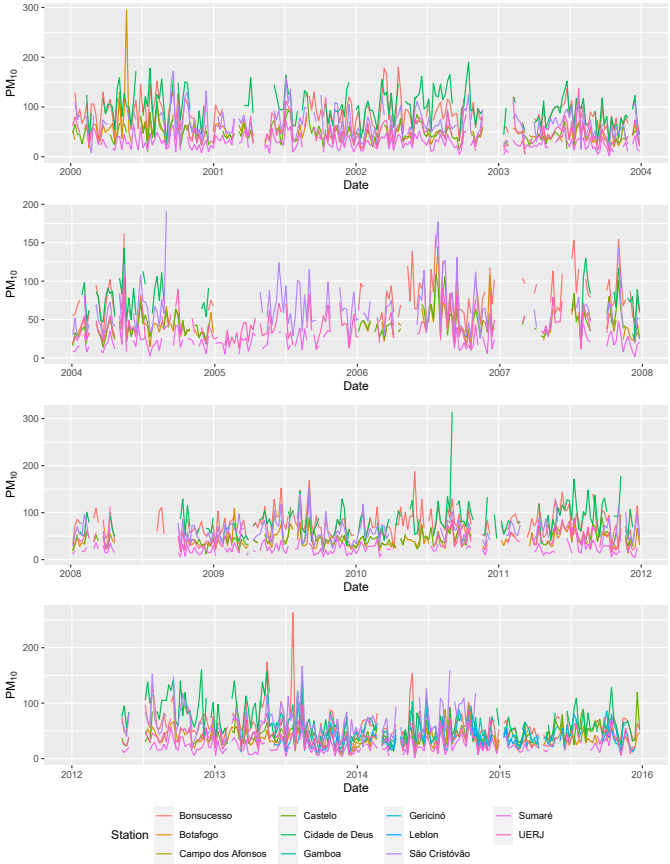


Figure 11: Time series of PM_{10} per monitoring station.

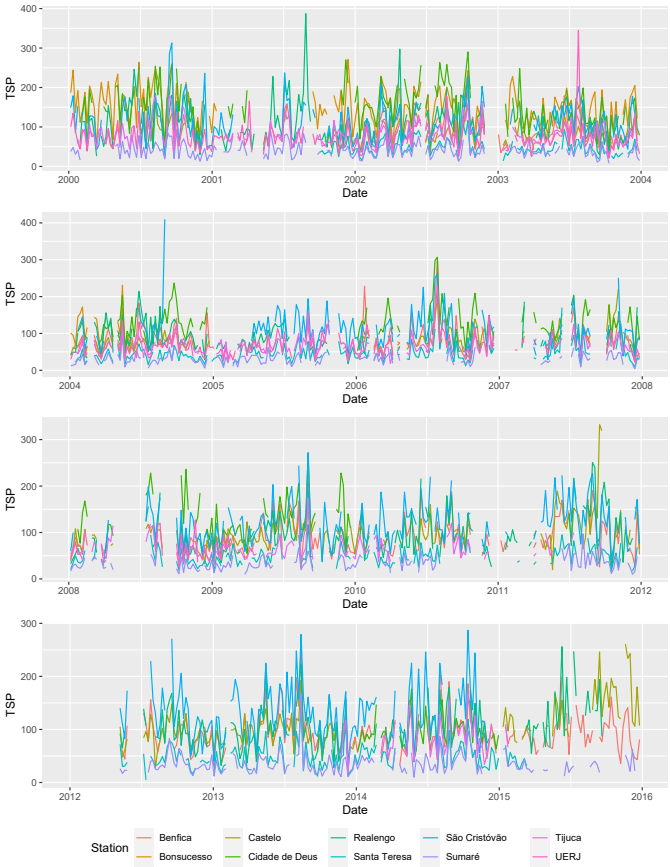


Figure 12: Time series of TSP per monitoring station.

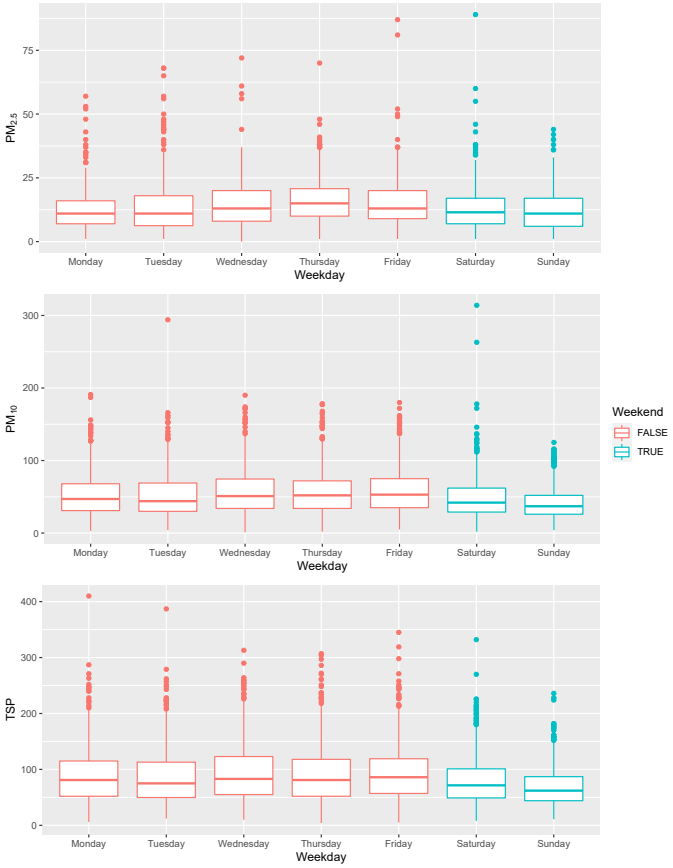


Figure 13: Boxplots of the air pollutant indexes per monitoring station.

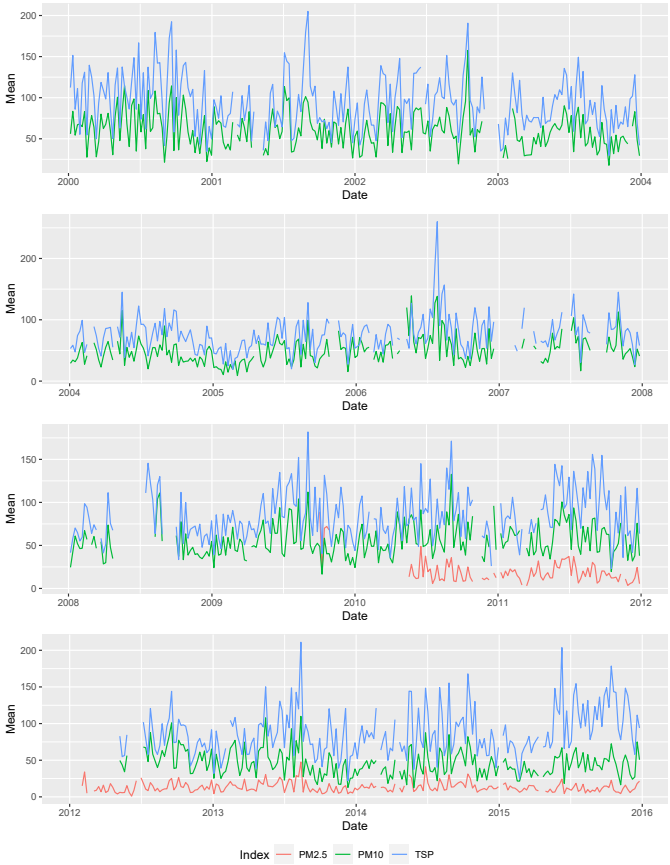


Figure 14: Time series of the air pollutant indexes aggregated in space by mean.

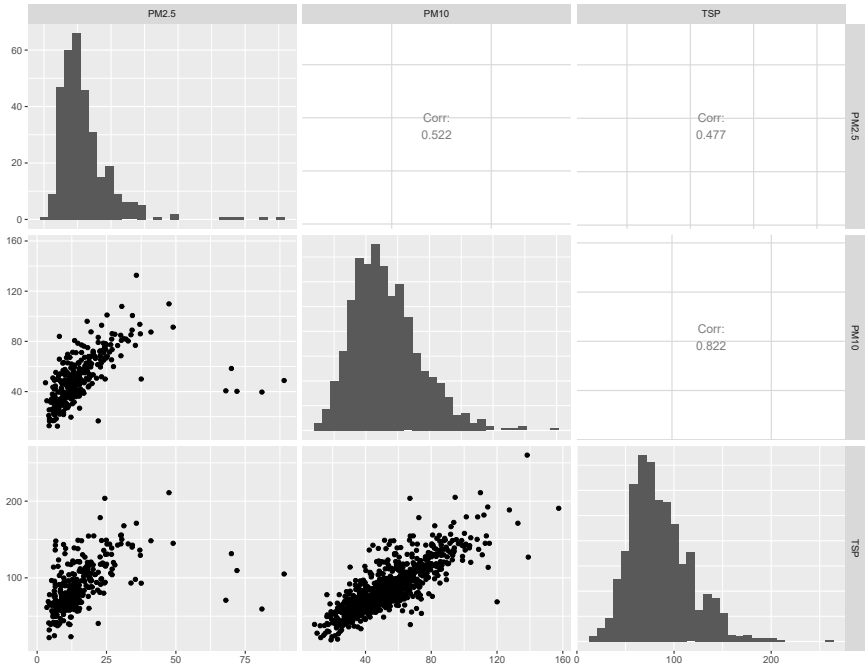
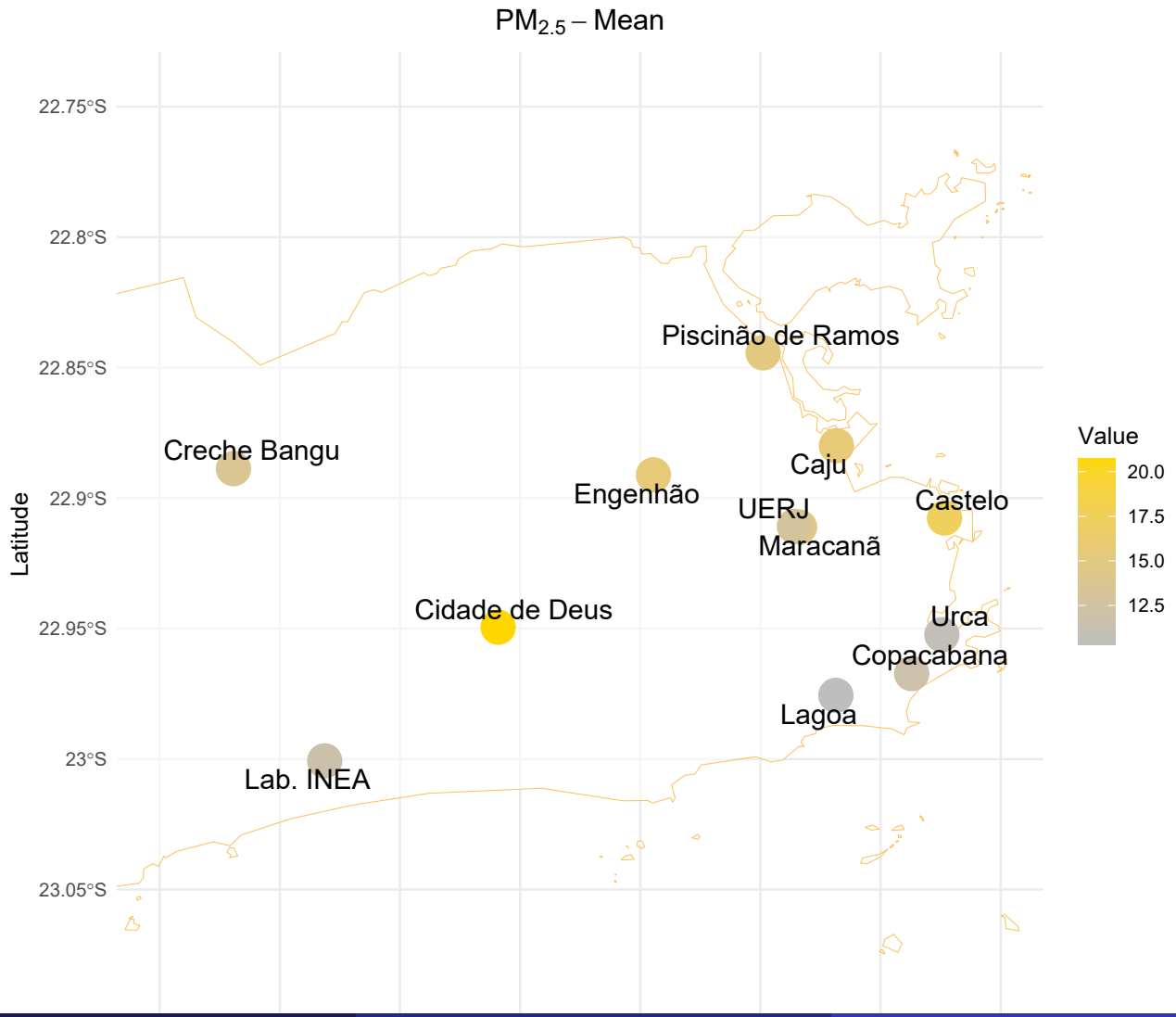
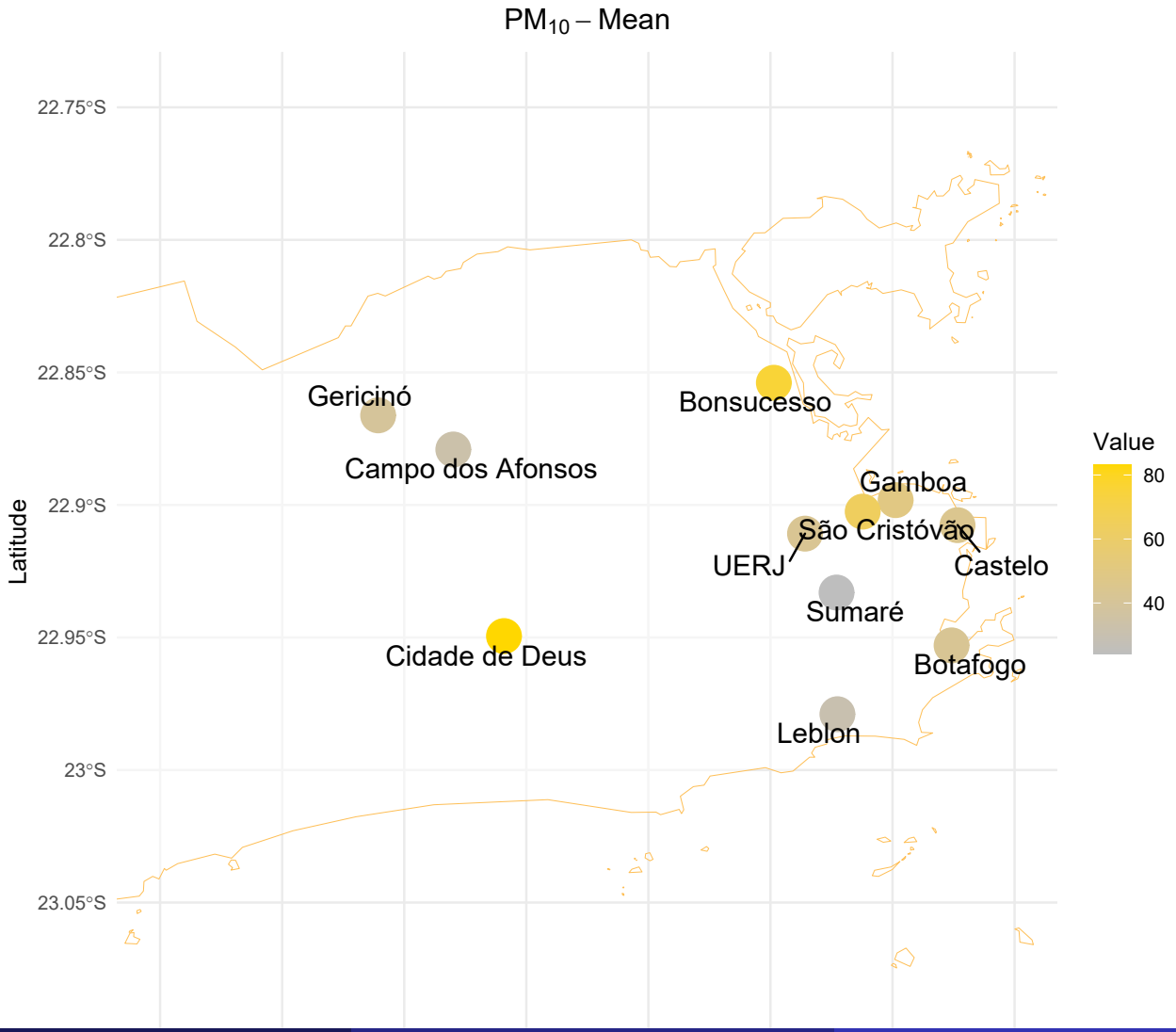
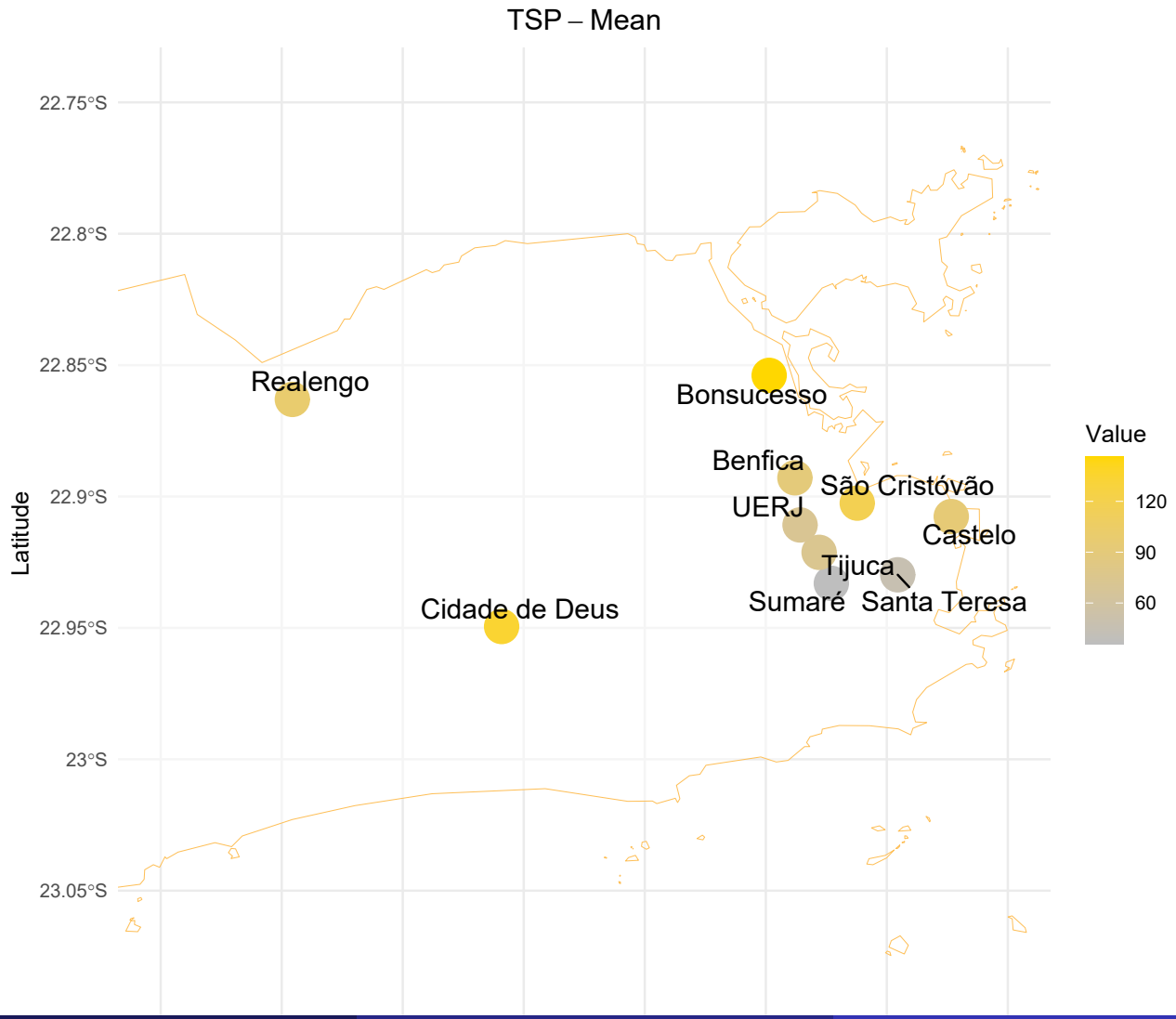


Figure 15: Relationship among air quality indexes aggregated in space by mean.


PM2.5 – Mean
22.75°S
22.8°S
22.85°S
22.9°S
22.95°S
23°S
23.05°S
Latitude
Piscinão de Ramos
Creche Bangu
Caju
Engenhão
UERJ
Castelo
Maracanã
Cidade de Deus
Urca
Copacabana
Lagoa
Lab. INEA
Value
20.0
17.5
15.0
12.5


PM10 – Mean
22.75°S
22.8°S
22.85°S
22.9°S
22.95°S
23°S
23.05°S
Latitude
Gericinó
Campo dos Afonsos
Bonsucesso
Gamboa
São Cristóvão
UERJ
Castelo
Sumaré
Cidade de Deus
Botafogo
Leblon
Value
80
60
40


TSP – Mean
22.75°S
22.8°S
22.85°S
22.9°S
22.95°S
23°S
23.05°S
Latitude
Realengo
Bonsucesso
Benfica
São Cristóvão
UERJ
Castelo
Tijuca
Sumaré
Santa Teresa
Cidade de Deus
Value
120
90
60

Multivariate response model

Suppose we do not observe only one but q ($q > 1$) response variables in a discrete set of periods of time $t = 1, \ldots, T$ and set of locations in the continuous space $S = \{s_1, \ldots, s_N\}$.

In the matrix notation, we have:

$$
\begin{aligned}
Y_t &= F_{1t}\Theta_{1,t} + \nu_{1t}, & \nu_{1t} &\overset{iid}{\sim} N(0, V_1, \Sigma) \\
\Theta_{1,t} &= F_{2t}\Theta_{2,t} + \nu_{2t}, & \nu_{2t} &\overset{iid}{\sim} N(0, V_2, \Sigma), \\
\Theta_{2,t} &= G_t\Theta_{2,t-1} + w_t, & w_t &\overset{iid}{\sim} N(0, W, \Sigma)
\end{aligned}
$$

- The matrix of regression parameters $\Theta_{1,t}$ is $Np \times q$ and it can be written as a collection of blocks $\Theta_{1,t} = [\Theta_{1,1,t}, \ldots, \Theta_{N,1,t}]'$, where $\Theta_{j,1,t}$ is $p \times q$ such that $Np = N \times q$, $\Theta_{2,t}$ is $p \times q$, F_{1t} is $N \times Np$, F_{2t} is $Np \times q$, and G_t is $p \times p$

- The matrices F_{1t} , F_{2t} and G_t are assumed known, with F_{1t} and F_{2t} incorporating covariate values.

- We do not impose spatial structure in ν_{1t}, for $t = 1, \ldots, T$, and define $V_1 = I_N$. The variance matrix V_2 is specified through parametric structures describing spatial dependency.

- We can define $V_2 = CV$ which corresponds to $\nu_{2t} \sim N(0, V\rho(\phi, \cdot), \Sigma)$, where C is the matrix specified through the correlation function $\rho(\phi, \cdot)$ - which can be isotropic in a deformed space.

- Note that the imposed restriction states that part of the conditional covariance between elements of Y_t, which account for the different responses, is specified in Σ.

- The same restriction is imposed to the conditional covariances of $\Theta_{1,t}$ and $\Theta_{2,t}$, with all sharing the same structure Σ to account for the covariance between elements coming from different responses.

- This assumption is usually made in multivariate dynamic models.

- It is reasonable since the magnitude of the covariance between elements at different hierarchical levels are not imposed to be the same.

- The models above could be specified and operated without this restriction but that would substantially increase parameter dimensionality.

- The model is completed with similar priors proposed under the univariate model
- Inverted Wishart prior distributions were proposed for the covariance matrices Σ;

Inference and computational methods

- The posterior distribution cannot be analytically solved
- Samples will be drawn through an efficient algorithm using Gibbs sampler based on the MCMC proposed by Landim and Gamerman (2000).
- $\{\Theta_1\}$,Σ,V, and W are sampled through their full conditionals and $(\Theta_{2,0}, \{\Theta_2\})$ are sampled through FFBS (forward filtering backward sampling—Frühwirth-Schnatter).

Under development

- We are now developing the MCMC algorithm to sample from the posterior distribution of the model parameters.
- Our aim is to apply the proposed model to the pollutants data-set.
- We also intend to study methods to approximate the posterior distribution and overpass the computational burden of the MCMC, like Variational Bayes.

References I

Cabral, C. R. B., Lachos, V. H., and Madruga, M. R. (2012).
Bayesian analysis of skew-normal independent linear mixed models with heterogeneity in the random-effects population.
Journal of Statistical Planning and Inference, 142(1):181–200.

Castro Morales, F., Gamerman, D., and Paez, M. (2013).
State space models with spatial deformation.
Environmental and Ecological Statistics, 20(2).

De Bastiani, F., de Aquino Cysneiros, A. M., Uribe-Opazo, M., and Galea, M. (2015).
Influence diagnostics in elliptical spatial linear models.
TEST: An Official Journal of the Spanish Society of Statistics and Operations Research, 24(2):322–340.

References II

Lenfest Jr, L. W. (1986).
Ground-water levels and use of water for irrigation in the Saratoga Valley, southcentral Wyoming, 1980-81.
Water-Resources Investigation Report 84-4040, pages 84–4040.

Sampson, P. D. and Guttorp, P. (1992).
Nonparametric Estimation of Nonstationary Spatial Covariance Structure.
Journal of the American Statistical Association, 87(417):108–119.

Schmidt, A. M. and O'Hagan, A. (2003).
Bayesian Inference for Nonstationary Spatial Covariance Structures via Spatial Deformations.
Journal of the Royal Statistical Society Series B, 65:743–758.

References III

Yan, J., Cowles, M. K., Wang, S., and Armstrong, M. P. (2007). Parallelizing MCMC for Bayesian Spatiotemporal Geostatistical Models.
Statistics and Computing, 17(4):323–335.