Spatio-temporal processes,

including dynamic linear models

Dani Gamerman

Departamento de Métodos Estatísticos (DME)

Instituto de Matemática - UFRJ

dani@im.ufrj.br

PASI on Spatio-temporal Statistics

Búzios, 19 June 2014

Content

Presentation based on GAMERMAN, D. (2010). Dynamic Spatial Models, Including Spatial Time Series. Chapter 24 of Handbook of Spatial Statistics (Eds.: A. E. Gelfand, A. E., Diggle, P., Guttorp, P. e Fuentes, M.). London: Chapman & Hall.

1. Introduction

Consider a process $X(s,t)$ varying with $s \in S \subset R^d$ and $t \in T \subset R$.

I will be concentrating on $d = 2$ and $T = \{t_1, ..., t_n\}$

[Theoretically, continuous T simply means that $d \rightarrow d+1$]

2 routes are usually taken for modeling:

- direct (observation-driven)
- latent (parameter-driven)

Routes are somewhat disjoint

2. Direct approach

Covariance function (CF): $Cov[X(s,t), X(s',t')] = K(s,s';t,t')$

Too general for reliable estimation without massive replications over $S \times T$

Simplifying assumptions:

1) Temporal stationarity: $K(s,s';t,t') = f(s,s';t-t')$

2) Spatial stationarity: $K(s,s';t,t') = f(s-s';t,t')$

3) Spatio-temporal stationarity: $K(s,s';t,t') = f(s-s';t-t')$

4) Separability: $K(s,s';t,t') = f_1(s,s')f_2(t,t')$

Cressie and Huang (1999), Gneiting (2002), many others:

stationary, non-separable CF's

3. Latent approach

Explain X behaviour through model components

CF's appear as a consequence

Knowledge on how X varies over (s,t), through known functions $\eta_i(s,t)$,

$\Rightarrow \; X(s,t) = \sum_i \beta_i \eta_i(s,t) \quad (+e(s,t) \;)$

errors $e(s,t)$ do not carry any spatio-temporal correlation

Usually not feasible $\longrightarrow$ knowledge rarely exists

Even when it does, there may remain substantial correlation

Idea can be used to span the space of possible representations

$\Rightarrow \; X(s,t) = \sum_i \beta_i \eta_i(s,t) \quad (\, +e(s,t) \,)$

Now: unknown η_i functions form a basis for (smooth) functions of $S \times T$

These functions adequately chosen to cover the entire space $S \times T$

This may require too many η_i's

Alternative: tensor product

$\{\eta_i(s,t)\}$ replaced by $\{\phi_i(s)\}$ and $\{\alpha_j(t)\}$

$\Rightarrow \; X(s,t) = \sum_i \phi_i(s) \alpha_i(t) \quad (\, +e(s,t) \,)$

Substantial dimension reduction, but there is a cost

More importantly, relevant correlation may still remain

General approach: allow $\{\eta_i(s,t)\}$ to be stochastic

How? to be detailed later

Simplified approach: allow $\{\phi_i(s)\}$ and/or $\{\alpha_j(t)\}$ to be stochastic

$\Rightarrow \ X(s,t) = \sum_i \phi_i(s)\alpha_i(t) \quad (\ +e(s,t) \ )$

$\{\phi_i(s)\}$ are Gaussian processes (GP) and/or $\{\alpha_j(t)\}$ are time series

Usual time series model: autoregressive

Example - AR(1): $\alpha(t) = G(t)\alpha(t-1) + w(t), \quad w(t) \quad iid \quad N(0,\Sigma)$

Lots of references: Mardia, Goodall, Redfern & Alonso (1998), Wikle & Cressie (1999), Stroud, Muller & Sansó (2001), Calder (2007), Sansó, Schmidt & Nobre (2008), ...

Also leads to non-separable CF's

4. Dynamic Gaussian processes

Set $\eta(s,t) = (\eta_1(s,t), ..., \eta_m(s,t))$

DGP: $\eta(s,t) = G(t)\ \eta(s,t-1) + w(s,t), \quad w(\cdot,t) \quad iid \quad {}_mGP(0,\rho)$

Process is completed with initialization $\eta(s,1) \sim {}_mGP$

Leads to non-separable but now also temporally non-stationary CF's

Simplest DGP: univariate spatio-temporal random walk

$$\eta(s,t) = \eta(s,t-1) + w(s,t), \quad w(\cdot,t) \quad iid \quad GP$$

Useful model for smooth spatio-temporal variation of the data $Y(s,t)$

$$Y(s,t) = \eta(s,t) + \ WN \ \ error$$

Can handle spatio-temporal heterogeneity of other model components

4.1. Regression

Assume the presence of covariates $Z(s,t)$ associated with data $Y(s,t)$

Standard approach: $Y(s,t) = Z(s,t)^T\eta + X_0(s,t) + WN \;\; error$

with $X_0 \sim DGP$

Spatio-temporal heterogeneity may also be present in the regression part

Revised approach: $Y(s,t) = Z(s,t)^T\eta(s,t) + X_0(s,t) + WN \;\; error$

with $X = [\eta, X_0] \sim DGP$

The mean of the DGP may be separated from $X(s,t)$

Now, $X(s,t) = X(t) + X^*(s,t)$ where X^* is a zero-mean DGP

Model completed with an evolution for time-varying mean $X(t)$

4.2. Trend

Assume the DGP $X(s,t)$ is subject to spatio-temporal variations

These are modeled with an auxiliary growth process $\gamma(s,t)$

$$\begin{aligned} X(s,t) &= X(s,t-1) + \gamma(s,t) + w_X(s,t), \quad w_X(\cdot,t) \quad iid \quad GP \\ \gamma(s,t) &= \gamma(s,t-1) + w_\gamma(s,t), \quad w_\gamma(\cdot,t) \quad iid \quad GP \end{aligned}$$

This is in the form of a bivariate DGP with $G(t) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

Acceleration and higher order variations can be equally defined

Useful for estimation of trend features and specially for prediction

4.3. Seasonality

Assume the DGP $X(s,t)$ is subject to seasonal variations

Seasonality given by a sinusoidal wave (temperature): single harmonic

Requires an auxiliary process $\gamma(s,t)$

$$\begin{pmatrix} X(s,t) \\ \gamma(s,t) \end{pmatrix} = \begin{pmatrix} \cos(2\pi/q) & \sin(2\pi/q) \\ -\sin(2\pi/q) & \cos(2\pi/q) \end{pmatrix} \begin{pmatrix} X(s,t-1) \\ \gamma(s,t-1) \end{pmatrix} + w_S(s,t),$$

where $w_S(\cdot,t) \quad iid \quad {}_2GP$ and q is the seasonal cycle length.

More elaborate seasonal patterns $\longrightarrow$ additional bivariate processes

$(X_2, \gamma_2), \ldots, (X_{[q/2]}, \gamma_{[q/2]})$ associated with 2nd, ... , $[q/2]$th harmonics

Simpler form: $\sum_{j=1}^{q} X(s,t-j) = w(s,t)$, where $w(\cdot,t) \quad iid \quad {}_1GP$

but requires q-dimensional DGP

5. Applications

DGP can be used in models for 3 types of spatial data

Directly applied to the mean structure of spatio-temporal data $Y(s,t)$

5.1. Geostatistics - continuous space

Spatial version of generalized dynamic linear models

$$Y(s,t) \sim \mathcal{F} \text{ with mean } \mu(s,t) \quad \left[\text{eg } N\left(\mu(s,t), \sigma^2 \right) \right]$$

$$g[\mu(s,t)] = F(s,t)^T X(s,t) \quad [F^T = (Z^T, 1) \text{ and } X = (\eta, X_0)]$$

$$X(s,t) = G(t)X(s,t-1) + w_X(s,t), \quad w_X(\cdot,t) \quad iid \quad GP$$

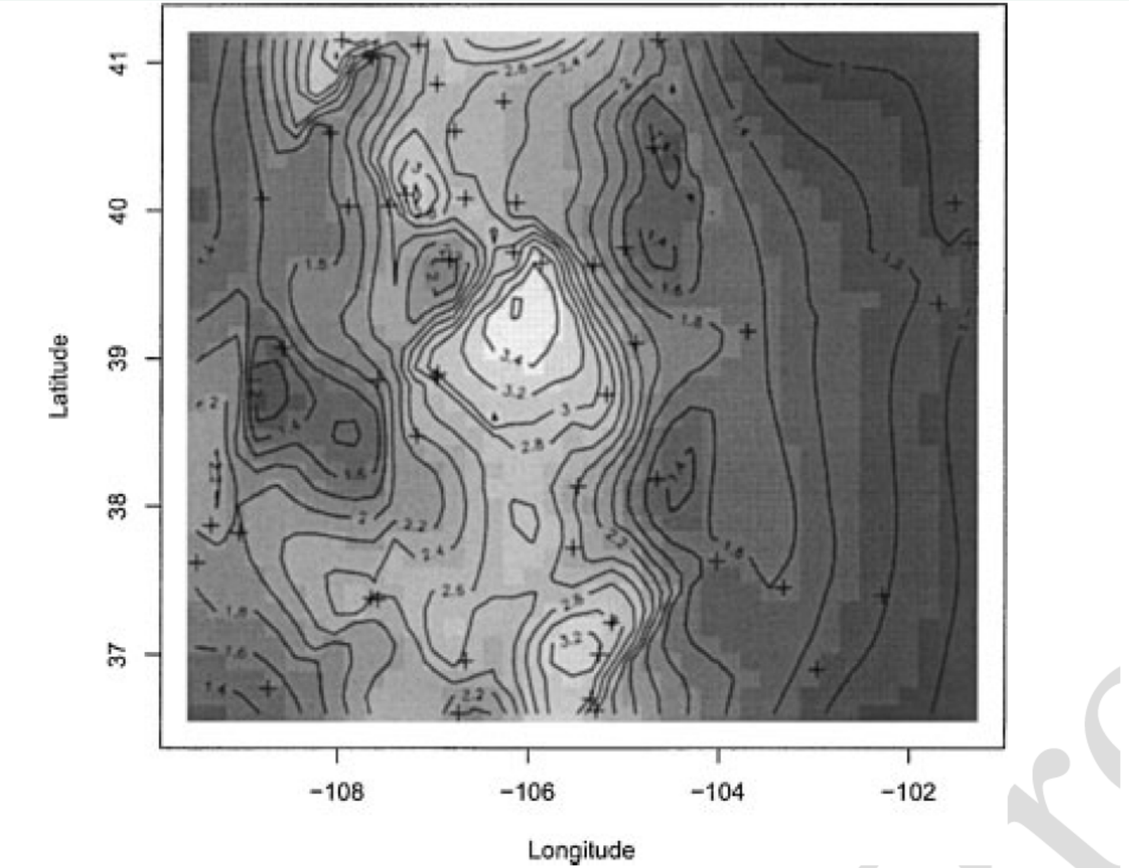
Illustration: Effect of precipitation (Prec) over temperature (Temp)

$$\begin{aligned}
\texttt{Temp}(s,t) &= \beta_0(s,t) + \beta_1(s,t)\ \texttt{Prec}(s,t) + v(s,t) \\
\beta_0(s,t) &= \beta_0(s,t-1) + w_0(s,t), \quad w_0(\cdot,t) \quad iid \quad GP \\
\beta_1(s,t) &= \beta_1(s,t-1) + w_1(s,t), \quad w_1(\cdot,t) \quad iid \quad GP
\end{aligned}$$

Intercept β_0 and regression coef β_1 are spatio-temporally varying

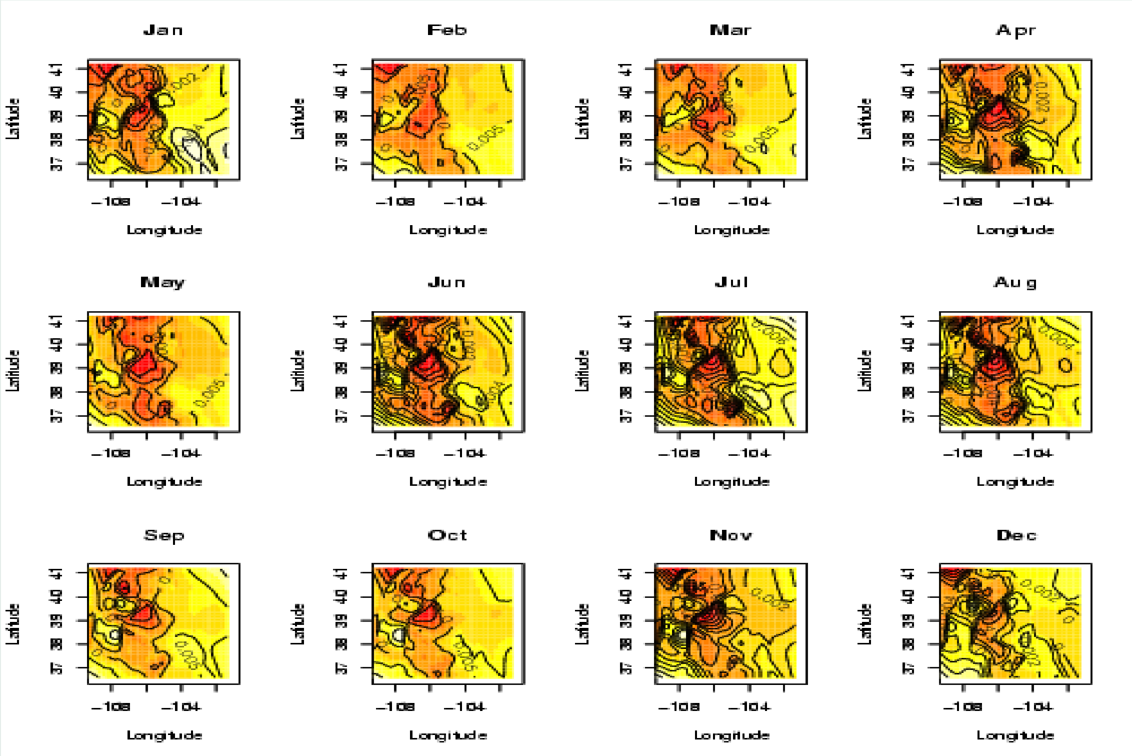
Disturbance processes w_0 and w_1 may be related

eg. linear transformations of independent GP's



S: region in the state of Colorado, USA ($+$ - monitoring stations)

T: $\{Jan/1997, \cdots, Dec/1997\}$



Posterior mean of the spatio-temporal variation of the regression coefficient β_1 for a region of the State of Colorado, USA (Gelfand, Banerjee and Gamerman, 2005).

5.2. Areal data

[based on Vivar and Ferreira (JCGS, 2012)]

Correlations are no longer based on CF's and Euclidean space

Mostly based on precision (inv. variance) matrices and neighbourhoods

Dynamic GP's replaced by dynamic Markov random fields (MRF)

Generalized spatio-temporal linear model

$$Y_{i,t} \sim \mathcal{F} \text{ with mean } \mu_{i,t} \quad \left[\text{ eg } N\left(\mu_{i,t}, \sigma^2 \right) \right]$$

$$g[\mu_{i,t}] = Z(s,t)^T X_{i,t}$$

$$X_t = G(t) X_{t-1} + w_t, \quad w_t \quad iid \quad MRF(0,Q) \equiv N(0, Q^{-1})$$

Q is sparse (full of 0's) ... $q_{ij} \neq 0$ indicates neighborhood (i,j)

Illustration:

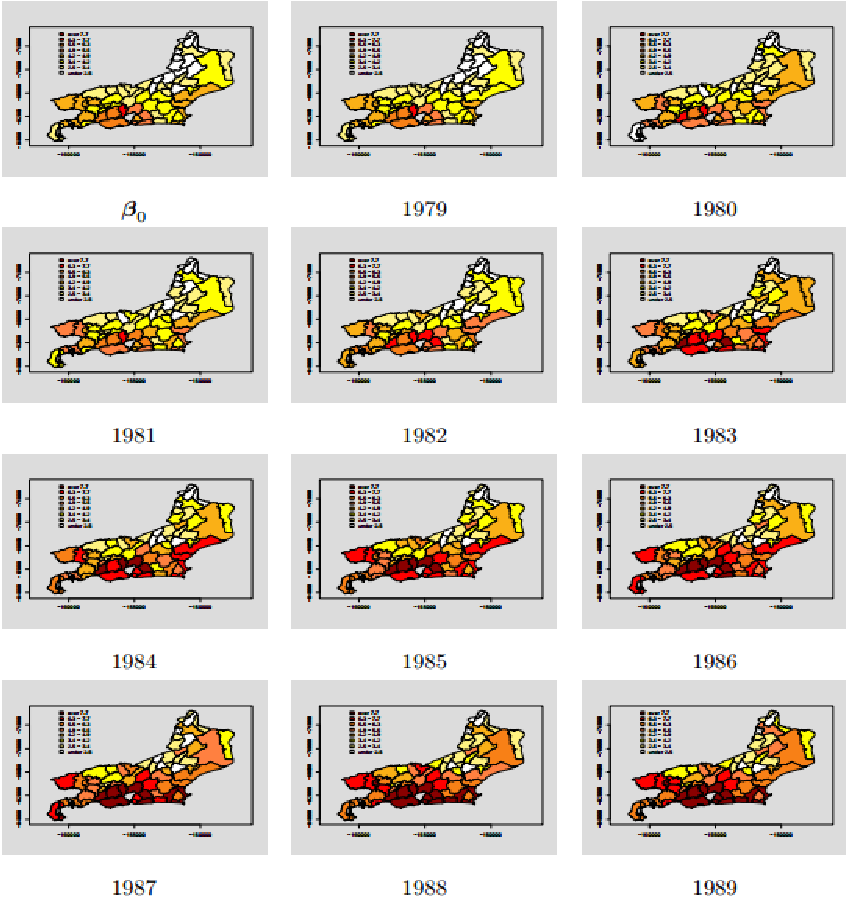
Evolution of the homicide rate in Rio de Janeiro state municipalities

Contamination model

$$
\begin{aligned}
Y_{i,t} &= \mu_{i,t} + e_{i,t}\ , \text{ where } e_{i,t} \sim N(0, \sigma^2_{i,t}) \\
\mu_t &= H\mu_{t-1} + w_t\ , \text{ where } v_t \sim N(0, Q^{-1}) \\
h_{i,j} &= c \times \begin{cases} 1, & \text{if } i = j \\ \alpha, & \text{if } i \text{ and } j \text{ are neighbors} \\ 0, & \text{otherwise} \end{cases}
\end{aligned}
$$

α is the contamination index

c is the contamination persistence



Posterior mean of the area level

5.3. Point pattern

based on Pinto Jr. (2014)

Data: space-time locations of occurrence of events

Usual models: PP, NHPP, Cox process (CP), log-Gaussian CP

LGCP (space only): $Y \sim NHPP(\lambda)$ and $\log \lambda = X \sim GP$

Liang et al (2008): $\log \lambda(s) = Z(s)^T \eta + \ X(s)\ ,\ X \sim GP$

Our extension:

$$
\begin{aligned}
Y \ &\sim \ PP(\lambda) \text{ where } \lambda : S \times T \to R^+ \\
g[\lambda(s,t)] \ &= \ Z(s,t)^T \eta(s,t) \qquad [\, eg. : \quad g = \log \,] \\
\eta(s,t) \ &= \ G(t)\eta(s,t-1) + w_\eta(s,t), \quad w_\eta(\cdot,t) \quad iid \quad GP
\end{aligned}
$$

Likelihood:

$$l(\lambda; Y) = \prod_i \lambda(s_i, t_i) \prod_t \exp\left[-\int_S \lambda(u, t)\, du\right]$$

difficult problem $\longrightarrow$ depends on an infinite-dimensional unknown function

Møller et al (1998) solution for CP: discretize λ over space $\longrightarrow$

problem: it is an approximation

Adams et al (2009) solution for CP: use thinning $\longrightarrow$

problems: does not extend easily for discrete time and for further model components and dimension explosion and MCMC...

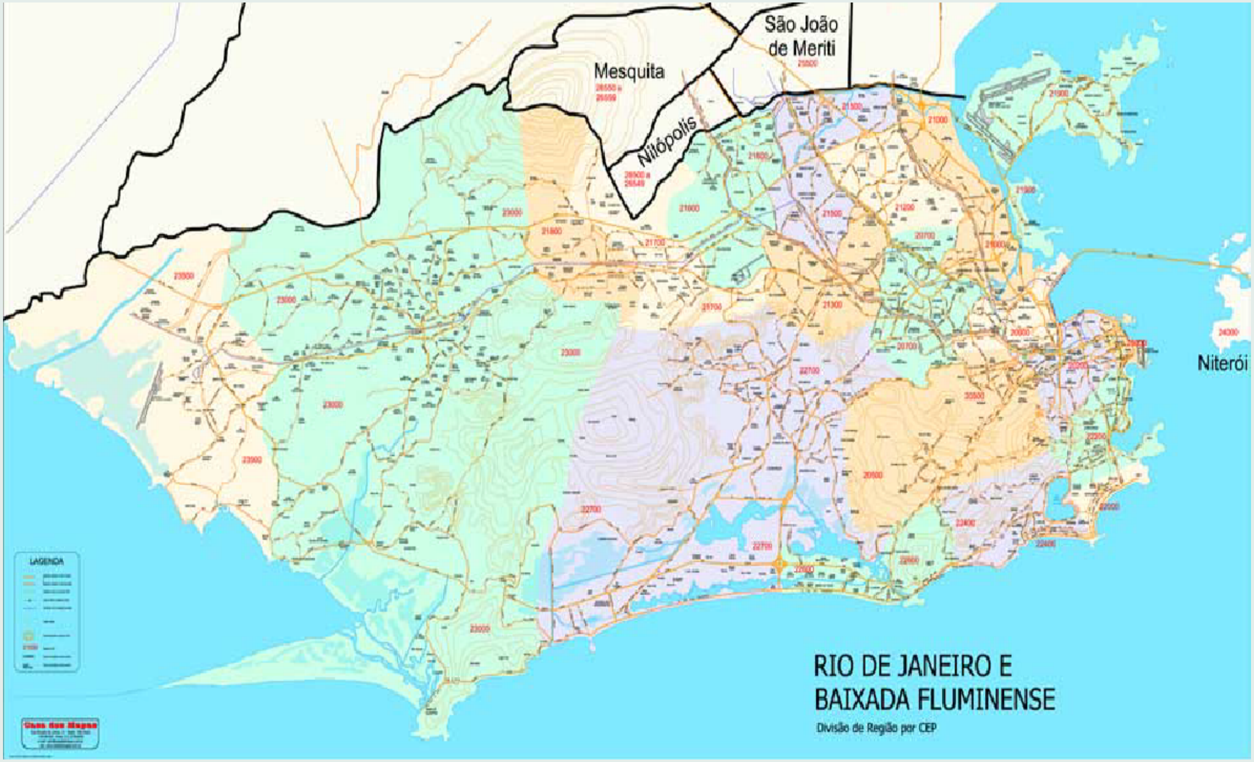
We are currently working on these solutions for our DGP model

Illustration:

Evolution of the vehicle casualties (theft/robbery) in Rio de Janeiro city

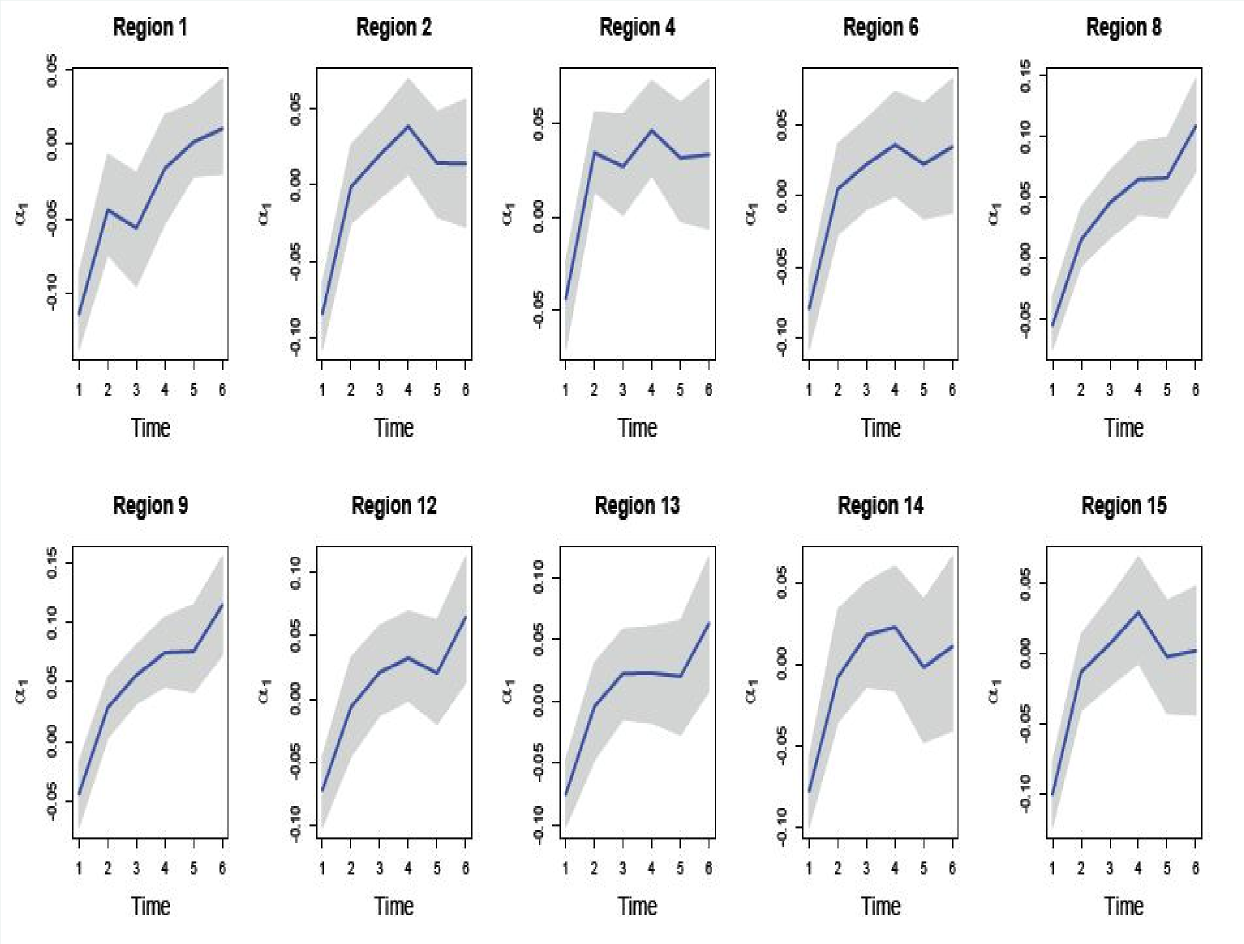
considering `Type` (private/commercial) and `Age` (manufacturing year)

$$
\begin{aligned}
Y &\sim PP(r\ \lambda) \text{ where } r, \lambda : S \times T \to R^+ \\
r(s,t;\texttt{Type},\texttt{Age}) &= \text{ total exposure at } (s,t) \quad [\texttt{known offset}] \\
\lambda(s,t;\texttt{Type},\texttt{Age}) &= \exp\{X_0(s,t) + \texttt{Type}\ X_1(s,t) + \texttt{Age}\ X_2(s,t)\} \\
\{X_0, X_1, X_2\}(s,t) &= \{X_0, X_1, X_2\}(s,t-1) + w_X(s,t),\ \ w_X(\cdot,t)\ iid\ GP
\end{aligned}
$$



S: Rio de Janeiro municipality

T: $\{sem1/2009, \cdots, sem2/2011\}$



Posterior mean of the spatio-temporal coefficient process of Age

6. Final comments

Flexible approach

Handles non-separable processes

Handles spatial non-stationary/stationary processes

Handles temporal non-stationary/stationary processes

More on the point process: Jony Pinto Jr. (here)

More on point process without discretization: F. B. Gonçalves (JSM 2014)

Thank you!

dani@im.ufrj.br