Point process with spatio-temporal heterogeneity

Jony Arrais Pinto Jr*

Universidade Federal Fluminense

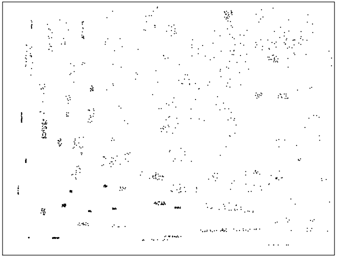
Universidade Federal do Rio de Janeiro

PASI
June 24, 2014

* - Joint work with Dani Gamerman and Marina S. Paez.

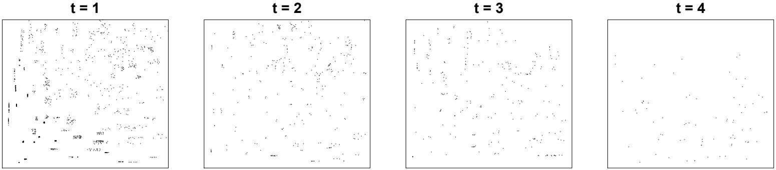
Summary

Introduction



- The event of interest is known.
- But the locations of occurrences of the event are unknown.
- Examples of events: infection of trees by a plague, deaths from stroke and vehicle casualties (theft/robbery) ...
- Geo-referenced data is very common: Ecology, Geography and Epidemiology.
- Point pattern is the set of these locations.

Introduction



• Point pattern is usually the result of a dynamic process that occurs both in space and in time.

• For example, event of interest is the infection of trees by a plague in a location.

• This process evolves over time as new trees are born and older trees die.

$\rightarrow$ spatial point pattern resulting from this process has a temporal dynamic.

Introduction

- Heterogeneity requires flexible models to capture this spatio-temporal variation.
- Geo-referenced data with the precise spatial location $\rightarrow$ study of point patterns.
- Connections with explanatory variables.
- Agronomist's interest: effects of spatial and unit-specific factors in the pattern of infection of trees by a plague and possible changes of these effects over time

$\rightarrow$ design plans of action to intervene where the infection of trees is larger.

Introduction

- Models for these types of data: spatio-temporal point processes.
- The literature for analyzing data with spatio-temporal heterogeneity is well developed.
- Brix and Diggle (2001): flexible class of spatio-temporal point process based on log-Gaussian Cox model.
- Diggle et al. (2005): model with deterministic spatial, temporal and spatio-temporal components.
- Reis et al. (2013): deterministic and stochastic component and included a dynamic structure in the temporal component.

Introduction

- They assumed the effects of the covariates, when they are considered, to be the same over space.
- May be appropriate in many practical situations but...
- Not a realistic assumption for dataset with a large spatio-temporal heterogeneity in the effect of some explanatory covariates.
- We would like to consider spatial and unit-specific covariates (Liang et al., 2009).
- **Purpose of this work**: models allowing spatio-temporal variation of the effects.

Spatial Point Patterns

- $X = \{X(\mathbf{s}) : \mathbf{s} \in S\}$, where $S \subseteq \Re^d, d > 0$ and

$$X(\mathbf{s}) = \begin{cases} 1, & \text{if the event of interest occurred in } \mathbf{s}, \\ 0, & \text{otherwise.} \end{cases}$$

- X can be unequivocally identified with occurrence set $\{\mathbf{s}_1, \ldots, \mathbf{s}_n\}$, $\mathbf{s}_i \in S$.

- Most common model: (non-homogeneous) Poisson process with intensity function $\Lambda(\cdot) = \{\Lambda(\mathbf{s}) : \mathbf{s} \in S\}$.

- Notation: $X \sim PP\,(\Lambda(\cdot))$.

- **[log-Gaussian] Cox process (LGCP)**: Λ_v are random [with $\log \Lambda_v \sim GP$].

Spatial Point Patterns

- $Y(\cdot)$, defined in S, is said to be isotropic Gaussian if $\forall\ n > 1$ and $\{s_1, \ldots, s_n\} \in S$

$$(Y(s_1), \ldots, Y(s_n))' \sim N(\mu \mathbf{1}, \tau^{-1}\mathbf{R}_\gamma), \tag{1}$$

denoted by

$$Y(\cdot) \sim GP(\mu, \tau, \rho_\gamma), \tag{2}$$

where $\mathbf{R}_\gamma$ is a correlation matrix with elements $R_{i,j} = \rho_\gamma(||\mathbf{s}_i - \mathbf{s}_j||)$.

- Spatio-temporal point pattern: $X(\cdot, \cdot)$ and $\Lambda(\cdot, \cdot)$.

- **Covariate information**: $(z_1(\mathbf{s}, t), \ldots, z_{p_1}(\mathbf{s}, t))'$ and $(v_1, \ldots, v_{p_2})'$.

- Model: continuous space and discrete time.

Likelihood

- Consider collections $\{X_v(\mathbf{s},t) : \mathbf{v} \in \mathcal{V}\}$ of Poisson point patterns and $\{\Lambda_v(\mathbf{s},t) : \mathbf{s} \in S, t \in \{1,\ldots,T\}$ and $\mathbf{v} \in \mathcal{V}\}$ of intensities, for covariate configuration $\mathbf{v}$.

- The likelihood is given by

$$L(\Lambda(\cdot,\cdot)) = \prod_{v\in\mathcal{V}} L(\Lambda_v(\cdot,\cdot)),$$

$$\text{where } L(\Lambda_v(\cdot,\cdot)) = \prod_{i=1}^{n_v} \Lambda_v(\mathbf{s}_{v,i}, t_{v,i}) \exp\left\{ -\sum_{t=1}^{T} \int_S \Lambda_v(\mathbf{s},t) d\mathbf{s}dt \right\}, \qquad (3)$$

n_v is the number of events observed for the configuration $\mathbf{v}$,
$\mathbf{s}_{v,i}$ is the location of the i^{th} event, for $i = 1, \ldots, n_v$,
$t_{v,i}$ is the time of the i^{th} event, for $i = 1, \ldots, n_v$.

Space-time varying coefficients model

$$
\begin{aligned}
X_v &\sim PP(\Lambda_v(\cdot,\cdot)), \forall\ v \in \mathcal{V},\\
\Lambda_v(\mathbf{s},t) &= r(\mathbf{s},t,v)\lambda(\mathbf{s},t,v), \forall \mathbf{s} \in S, \quad v \in \mathcal{V},\\
\log\lambda(\mathbf{s},t,v) &= \mathbf{z}(\mathbf{s},t)'\boldsymbol{\beta}(t) + \mathbf{v}'\boldsymbol{\alpha}(\mathbf{s},t) + w(\mathbf{s},t),\\
\boldsymbol{\beta}(t) &= \boldsymbol{\beta}(t-1) + \epsilon_\beta(t), \quad \epsilon_\beta(t) \sim N(\mathbf{0}, \Omega_t),\\
\boldsymbol{\alpha}_l(\mathbf{s},t) &= \boldsymbol{\alpha}_l(\mathbf{s},t-1) + \epsilon_{\alpha_l}(\mathbf{s},t), \quad \epsilon_{\alpha_l}(\mathbf{s},t) \sim PG\left(0, \tau_{\alpha_l}, \rho_{\gamma_{\alpha_l}}\right),\\
w(\mathbf{s},t) &= w(\mathbf{s},t-1) + \epsilon_w(\mathbf{s},t), \quad \epsilon_w(\mathbf{s},t) \sim PG\left(0, \tau_w, \rho_{\gamma_w}\right).
\end{aligned}
$$

- $\Lambda(\cdot,\cdot)$ (multiplicative decomposition), $r(\cdot,\cdot,\cdot)$ representing a known offset (required for standardization).
- Time-varying coefficients: $\boldsymbol{\beta}(\cdot)$.
- Space-time varying coefficients: $\boldsymbol{\alpha}_l(\cdot,\cdot)$ and $w(\cdot,\cdot)$.
- $\alpha(\cdot,\cdot)$ and $w(\cdot,\cdot)$ are stationary and isotropic GP in space and autoregressive and non-stationary in time.

Space-time varying coefficients model

$$
\begin{aligned}
X_v &\sim PP(\Lambda_v(\cdot,\cdot)), \forall\ v \in \mathcal{V}, \\
\Lambda_v(\mathbf{s}, t) &= r(\mathbf{s}, t, v)\lambda(\mathbf{s}, t, v), \forall \mathbf{s} \in S, \quad v \in \mathcal{V}, \\
\log \lambda(\mathbf{s}, t, v) &= \mathbf{z}(\mathbf{s}, t)'\boldsymbol{\beta}(t) + \mathbf{v}'\boldsymbol{\alpha}(\mathbf{s}, t) + w(\mathbf{s}, t), \\
\boldsymbol{\beta}(t) &= \boldsymbol{\beta}(t-1) + \epsilon_\beta(t), \quad \epsilon_\beta(t) \sim N(\mathbf{0}, \Omega_t), \\
\boldsymbol{\alpha}_l(\mathbf{s}, t) &= \boldsymbol{\alpha}_l(\mathbf{s}, t-1) + \epsilon_{\alpha_l}(\mathbf{s}, t), \quad \epsilon_{\alpha_l}(\mathbf{s}, t) \sim PG\left(0, \tau_{\alpha_l}, \rho_{\gamma_{\alpha_l}}\right), \\
w(\mathbf{s}, t) &= w(\mathbf{s}, t-1) + \epsilon_w(\mathbf{s}, t), \quad \epsilon_w(\mathbf{s}, t) \sim PG\left(0, \tau_w, \rho_{\gamma_w}\right).
\end{aligned}
$$

- Non-stationary $\beta(\cdot)$, $\boldsymbol{\alpha}_l(\cdot,\cdot)$ and $w(\cdot,\cdot)$ is one of the possibilities.
- The equations above define a generalization of log-Gaussian Cox process.
- Interactions between spatial and non-spatial covariates can be considered.

Discretizing log-Gaussian Cox processes

- The likelihood function for the model depends on uncountable functions $\boldsymbol{\beta}(\cdot)$, $\boldsymbol{\alpha}(\cdot,\cdot)$ and $w(\cdot,\cdot)$.
- This poses a difficult problem to handle.
- Exact solutions are only available in very limited cases and even then, they depend on a number of issues.
- Some of these issues are associated with the dimension of the number of occurrences, which is usually very large.

Discretizing log-Gaussian Cox processes

Reasonable option: approximations at the modeling level

Beněs et al (2002):

- S is partitioned into sub-regions $\{S_1, \ldots, S_N\}$.
- $r(\mathbf{s}, t, \mathbf{v}) = r_{k,t,v}$, $\boldsymbol{\alpha}(\mathbf{s}, t) = \boldsymbol{\alpha}_{[k,t]}$ and $\mathbf{w}(\mathbf{s}, t) = w_{[k,t]}, \forall \mathbf{s} \in S_k$.
- enforces homogeneity of the intensity rate within the sub-regions in time $t \rightarrow$ $\lambda(\mathbf{s}, t, \mathbf{v}) = \lambda_{k,t,v} = \exp\{\mathbf{v}_j' \boldsymbol{\alpha}_{[k,t]} + \mathbf{z}_{[k,t]}' \boldsymbol{\beta}_t + w_{[k,t]}\}$
- The integral in (3) becomes

$$\sum_{t=1}^{T} \int_S r(\mathbf{s}, t, \mathbf{v}) \lambda(\mathbf{s}, t, \mathbf{v}) d\mathbf{s} = \sum_{t=1}^{T} \sum_{k=1}^{N} r_{k,t,v} \lambda_{k,t,v} |S_k| \tag{4}$$

  $|S_k|$ is the volume of the kth sub-region, for $k = 1, \ldots, N$.

Intensity discretization also found in Møller et al (1998), Brix and Møller (2001) and Gamerman (1992), ...

Discretizing log-Gaussian Cox processes

- Waagepetersen (2004): posterior distributions of the intensities converge to the posterior distribution of the continuously-varying intensity when the volumes of the sub-regions tend to 0.

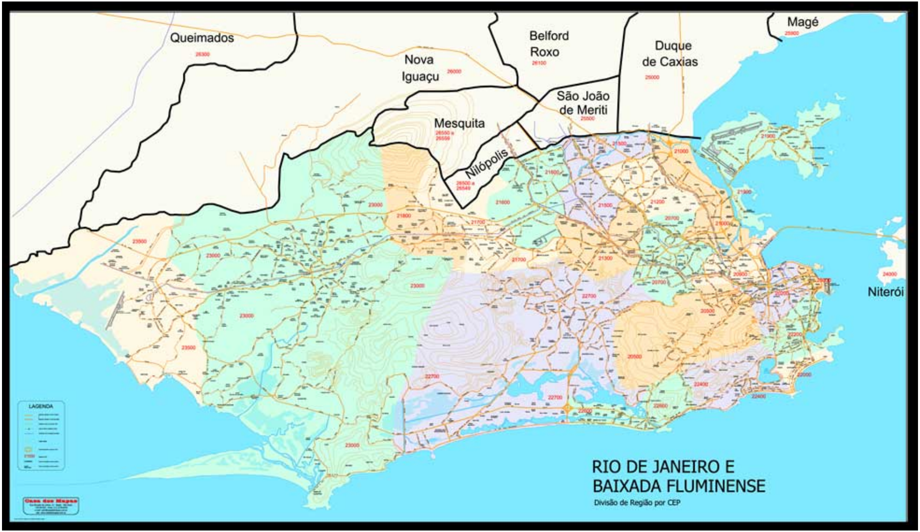
- If interest lies in the effect of a covariate at the region level rather than at a specific location, the discretization does not cause any limitation in the results.

- Number and sizes of the sub-regions must be appropriately chosen.

Data details - covariate information

- Data: The zip codes where cars covered by insurance have been stolen in Rio de Janeiro.
- 23,810 cars were stolen between 2009 and 2011.
- The time was discretized by semester (6 semesters).
- Unit-specific covariates v are:
  - v_1, manufacturing year;
  - v_2, car type (1, for private car and 0, commercial);
- $\mathbf{v} = (v_1, v_2)'$ is the vector of non-spatial covariates and the number of different configurations of these variables.
- $\#\mathcal{V} = 22$.

Discretized space



Effects varying over space and time

- Locations in Rio de Janeiro reflect different socioeconomic backgrounds.
- Great incentives to buy cars were experienced by the population in recent years.
- Theft pattern may be affected by this variation $\rightarrow$ effects should be as flexible as possible.
- Example: is passenger car thefts decreasing in the wealthiest areas of Rio de Janeiro over the last semesters?
- This question, for example, can only be answered by allowing interaction among the effects of covariates, space and time.

Prior distributions

- Mean and precision:
  - $\mu_x \sim N(0, 100), x = \alpha_1, \alpha_2, w;$
  - $\tau_x \sim G(1, 0.01), x = \alpha_1, \alpha_2, w;$

- GP correlation functions $\rho(||\mathbf{s}_i - \mathbf{s}_j||; \gamma) = \exp\{-||\mathbf{s}_i - \mathbf{s}_j||/\gamma\}$.

- Fonseca and Steel (2011): $\gamma_x \sim G(1, 0.3/\text{med}(d_{\mathbf{S}}))$,
  $\text{med}(d_{\mathbf{S}})$ = median of the distances among the 21 regions.

$\rightarrow$ weak identification of the range parameters γ.

- This well-known difficulty of spatial models, more pronounced here.

- Liang et al. (2009): fix the ranges at the median of the observed distances

- No significant changes for likelihood parameters but stabilized results for the hyperparameters.

Offsets

- The offsets $r_{k,t,v}$ were taken as populational size

$$r_{k,t,v} = \sum_{i=1}^{N_{k,v}} \frac{\#\text{days of annual policy} \quad i_{v,k,t}}{\#\text{days of the year}} \tag{5}$$

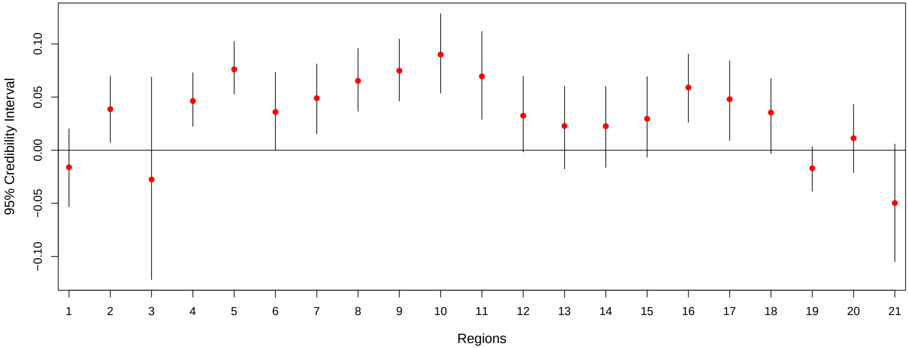
- $N_{k,v}$ is the number of cars with configuration v in region k.

- Requires knowledge of the population sizes of all configurations in each region for all periods of time.

Results

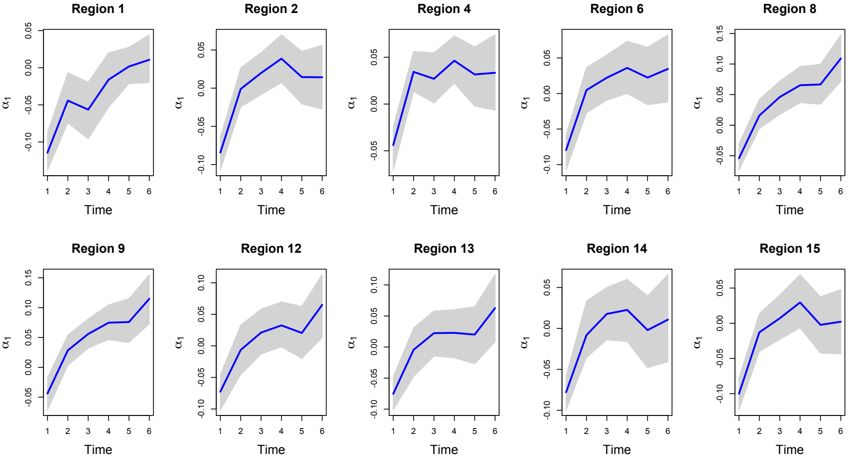
- Results below based on the 21 regions partition.
- Results were obtained via MCMC methods, with Winbugs.
- Convergence was ascertained by using 2 chains with different starting values.
- Correlation between successive chain draws was alleviate by thinning at every 100 iterations, after a burn-in period of 5,000 draws.
- The resulting sample consisted of 2,000 draws.

Manufacturing year effect: posterior median and CI(95%) of coefficient in $t = 4$



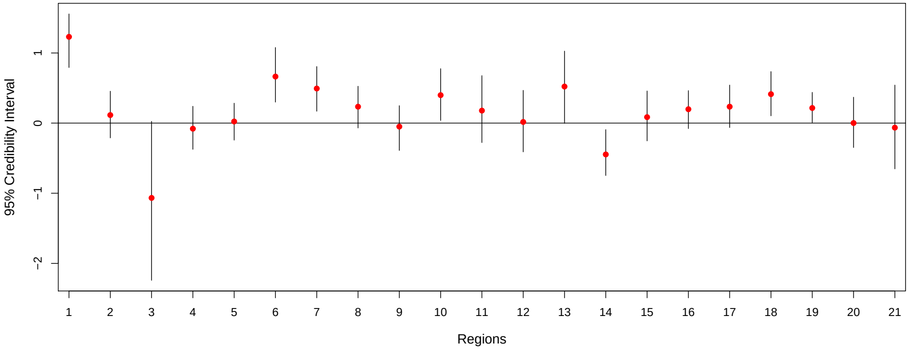
- Positive coefficients indicate a greater intensity of thefts of new cars.
- Negative coefficients are associated with the western region of Rio de Janeiro.
- The largest range of the CI is observed at region 3 (scarcity of information - small island).

Manufacturing year effect over time for some regions



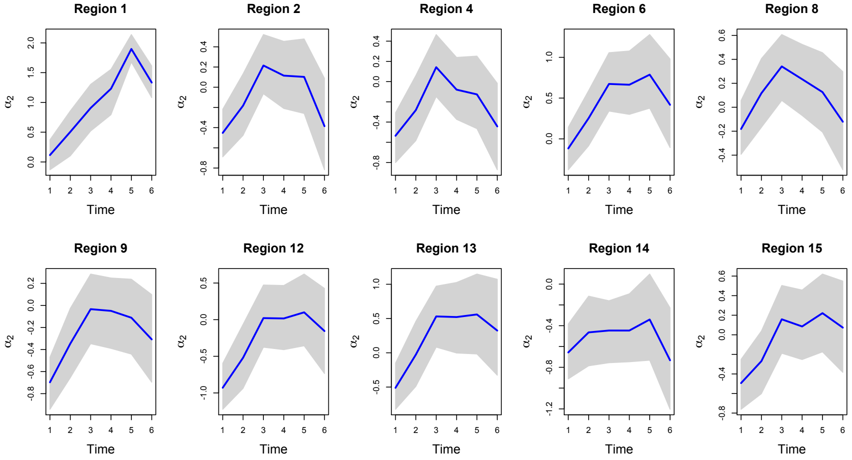
- In general, there is an increase of coefficients over time.

Type car effect: posterior median and CI(95%) of coefficient in $t = 4$



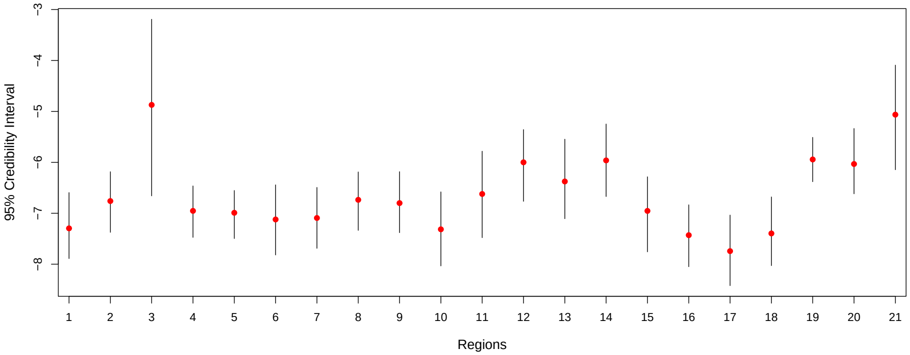
- Positive coefficients indicate a greater intensity of thefts of private cars.
- Largest coefficient: downtown.
- Negative coefficients: islands (3 and 14).

Type car effect over time for some regions



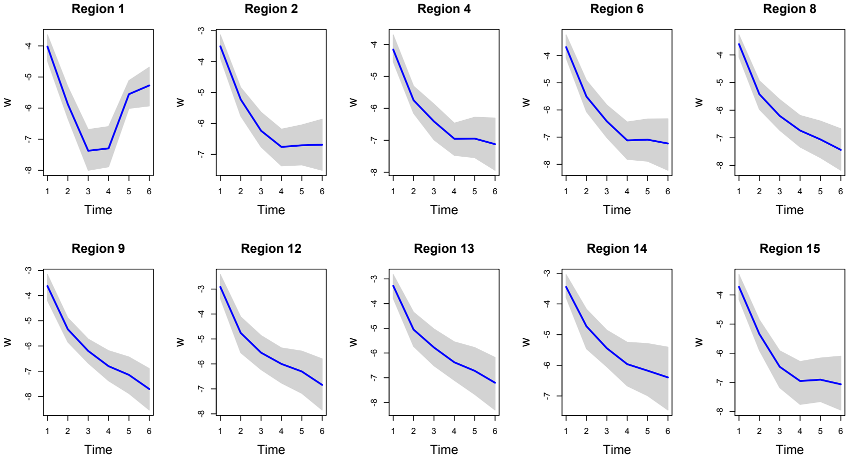
• For some regions, there is a significant growth in earlier periods, and a decrease in the final period.

Random intercept: posterior median and CI(95%) of coefficient in $t = 4$



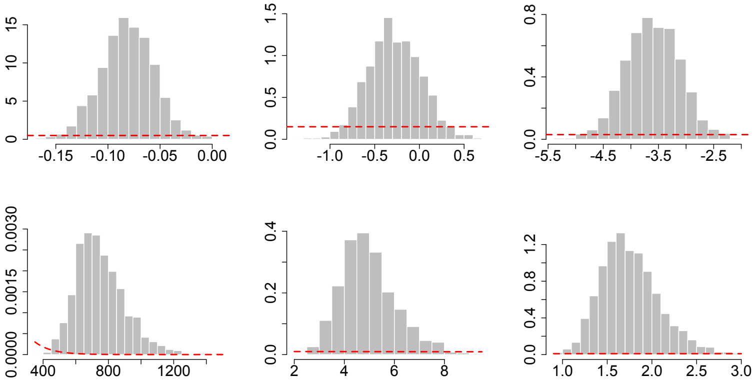
- w is responsible for capturing remaining spatio-temporal variation, not captured by the covariates.
- w has to compete with covariates for spatial and time variation in the model.
- Spatial variation of w is still significant.

Random intercept over time for some regions



- Temporal variation of w is still significant.

Posterior histograms for the GP means e precisions



- Top row - μ: year of manufacture, car type and w;
- Bottom row - τ: year of manufacture, car type and w;
- The dashed lines indicate the vague prior densities used.

Conclusions

- This work: hierarchical formulation to handle point patterns subject to the effect of covariates with large spatio-temporal heterogeneity.
- Heterogeneity spatio-temporal: state space and isotropic GP's tools $\rightarrow$ smooth variation of coefficients.
- Covariates have a strong spatio-temporal variation.
- Results showed the model can capture the spatial and time trend of the effects.
- Meaningful, concentrated posteriors were obtained with vague priors (exception range parameter).
- Finally, analysis of point patterns could benefit from a model without discretization.
- This is currently an active area of research (eg, Goncalves et al (2014)).

Thank you!

jarrais@est.uff.br