Statistical inference for spatio-temporal point process models

Thomas Opitz

BioSP, INRA
Avignon, France

Séminaire RESSTE – Processus de points
December 6, 2017

Program for this talk

- present **space-time (ST) point process modeling and inference**

  ⇒ focus on **fundamental concepts** and **most prominent model classes**

- point towards **software implementations** : R packages

- develop an application example : space-time modeling of **forest fires**

Notations

- S space, T time, ST space-time
- D_{ST} space-time support of the point process
- $|A|$ volume of set A
- $X = \{x_1, \dots, x_N\}$ a space-time point pattern with $x_i = (s_i, t_i)$
- N the (random) number of points in a point process
- $N(A)$ the random number of points in a space-time set $A \subset D_{ST}$
- $\lambda(s, t)$ space-time intensity
- $\lambda^{(k)}(y_1, \dots, y_k)$ kth order product density

Typical application contexts

(E) Empirical/correlative modeling vs. **(M) mechanistic modeling**

E-based approaches : "density-based", often regression-like, integrating covariates and random ST effects ⇒ ST intensity/risk mapping, ST covariance

M-based approaches : "individual-based" with direct interactions between "neighboring" points, causality

Often Markovian structures over time at density level (E) or point level (M)

Some application fields :

- biology/ecology : population dynamics (M), species distribution modeling (E), animal trajectories (M)
- epidemiology (E+M)
- environmental risks : wildfires (E), earthquakes (E+M)
- medicine/neurosciences : spike trains, ... (M)

What distinguishes ST from S ?

- "space-time" is 3D space, but with **oriented time axis**

  $\rightsquigarrow$ time asymmetry, dynamic and evolutionary models, causality ...

- **higher dimensionality of models and data** in ST

- **non separability** of S and T ?
  - if separable, may estimate S and T components separately
  - if nonseparable, **model building / estimation / selection may be challenging**

- how do we observe **arrival times** ?
  - observation in (quasi-)continuous time
    $\Rightarrow$ points may **arrive sequentially** $t_1 < t_2 < \ldots$ (**orderly process**)
    $\Rightarrow$ can use M-model,
    by specifying conditional intensity $\lambda(s, t \mid X_{t_i < t})$ (= hazard rate)
  - observation in discrete time
    $\Rightarrow$ may simplify E-modeling, but more complicated for M-modeling
  - arrival times may be not observed at all,
    but modeling of temporal dynamics may still be important

A little reminder for caution

Mechanisms generating a ST point pattern may be highly complex, BUT

- **complex models** are typically **difficult to simulate/estimate/validate**,
- **information** contained in the observed point pattern is often **quite limited.**

> *From a statistical perspective, a simple but well-identified model may be more valuable than an over-complicated model incorporating more parameters than can reasonably be estimated from the available data.*

[Diggle, 2006]

> *With four parameters I can fit an elephant, and with five I can make him wiggle his trunk.*

(attributed to John von Neumann)

Wei. *Least Square Fitting of an Elephant*. CHEMTECH, 5(2), 1975.
Mayer et al. *Drawing an elephant with four complex parameters*. Am. J. Phys. 78, 2010.
Williams. *Data fitting and elephants*. 2016

⚠ **Careful not to be overly ambitious with respect to model complexity.**

Intensity and interaction between points

ST intensity function $\lambda(s,t)$ characterizes expected point count for arbitrary sets A :

$$\mathbb{E}N(A) = \int_A \lambda(s,t)\,\mathrm{d}(s,t), \qquad A \subset D_{\mathrm{ST}}$$

$=$ first moments

ST interaction is about co-occurrence patterns of points,
e.g. K-function gives expected number of points in a ST buffer around arbitrarily chosen x_i

Notion of **separability** can apply to intensity and/or interaction :

intensity	$\lambda(s,t) = \lambda_S(s) \times \lambda_T(t)$
pairwise interaction	$K_{ST}(s,t) \propto K_S(s) K_T(t)$
conditional intensities	$\lambda(s,t \mid X_{t_i<t}) = \lambda_S(s \mid X_{t_i<t}) \times \lambda_T(t \mid X_{t_i<t})$
...	...

Modeling the space-time intensity function

Primary interest may lie in estimating the **space-time intensity function** $\lambda(s,t)$ where

$$\mathbb{E}\, N(A) = \int_A \lambda(s,t)\mathrm{d}(s,t).$$

Often, $\lambda(s,t)$ is **driven by covariates** (including S and T),
and a **generalized additive structure** may be useful :

$$\log \lambda(s,t) = \sum_i^k \beta_i z_i(s,t) + \text{nonlinear covariate effects} + \ldots$$

Disregard interaction between points $\Rightarrow$ fit a Poisson process $\sim \lambda(s,t)$ with likelihood

$$\exp\left(-\int_{D_{ST}} \lambda(x)\,\mathrm{d}x\right) \sum_{i=1}^{N} \lambda(x_i).$$

By (finely) **discretizing ST**, can estimate parameters in a GAM for **Poisson counts**.

Estimation of β_i is consistent under mild conditions
[Waagepetersen, 2007, Waagepetersen and Guan, 2009].

Model classes : Models with Gibbs specification

Gibbs specification with potential functions u_k : density of X has form

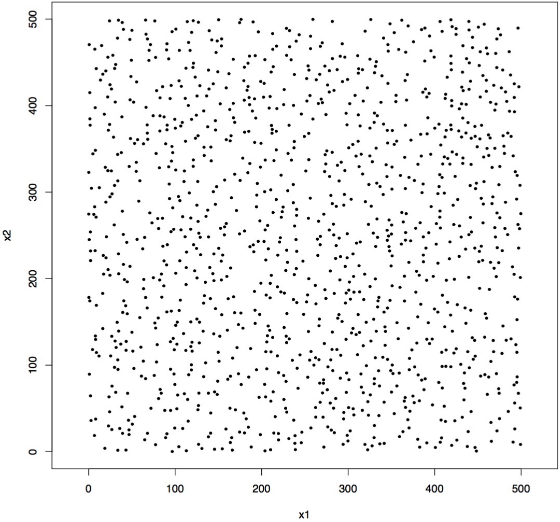
$$f(X) = C \exp\left(-\sum_{k=1}^{\infty}\sum_{i_1,\ldots,i_k} u_k(x_{i_1},\ldots,x_{i_k})\right), \quad C > 0$$

relative to a unit rate Poisson process

- mechanistic construction and interpretation
- Poisson process if $u_k \equiv 0$ for $k > 1$
- Gibbs models are **typical for interpoint repulsion**,
  but also some models for attraction (area interaction, Geyer, ...)
- **Markov point processes** [Van Lieshout, 2000, Van Lieshout, 2006]
  especially pairwise interaction processes with $u_k \equiv 0$ for $k > 2$
  $\Rightarrow$ conditional intensity of point in $x = (s,t)$ depends only on "neighbors" $x_i \sim x$
- ⚠ unknown parameter-dependent constant C
- ⚠ moments ($\lambda(s,t)$, K-function, ...) usually not known in closed form
- a simple ST example : hard-core process
  $\Rightarrow$ pairwise interaction : $u(x_1,x_2) = \infty$ if $(s_2 - s_1, t_2 - t_1) \in A$ with $A \subset S \times T$

Example : Gibbs process with inhibition

Stationary S process : multiscale model
Strong repulsion for point-to-point distance 0-4, weak repulsion for 4-15, weak attraction for 15-30
Metropolis–Hastings MCMC simulation



Model classes : Models with Cox specification

Cox process = Poisson process with random intensity $\Lambda(s,t)$

- $\Lambda(s,t)$ is a **ST random field**
- Cox processes cannot model repulsion between points,
  but allow **capturing the effect of unobserved/unavailable covariates**
- **Bayesian interpretation :** Poisson process with nonparametric prior $\Lambda(s,t)$
- models with "mechanistic" flavor : **shot-noise Cox processes**

$$\Lambda(s,t) = \sum_{i=1}^{N^P} M_i k(s - s_i^P, t - t_i^P)$$

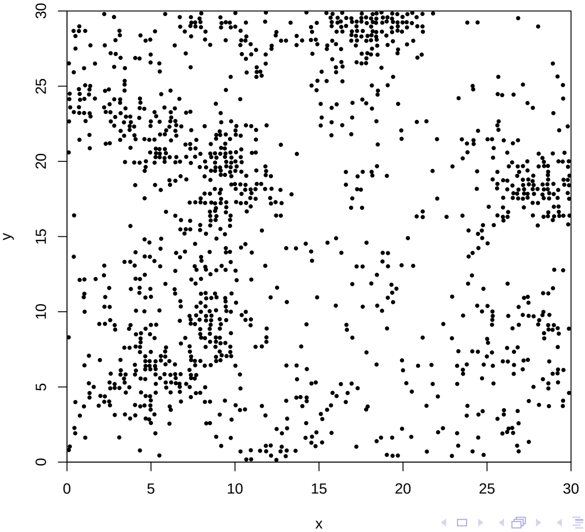
  with ST kernel k, parent Poisson points (s_i^P, t_i^P) and random "shots" $M_i > 0$
  $\rightsquigarrow$ **Neyman-Scott processes** if $M_i \equiv 1$
- **log-Gaussian Cox processes** are a popular choice for E-models :

$$\log \Lambda(s,t) = [\text{fixed covariate effects}] + [\text{random effects}]$$

  random effects = ST Gaussian random field such as $W_S(s)$, $W_T(t)$, $W_{ST}(s,t)$

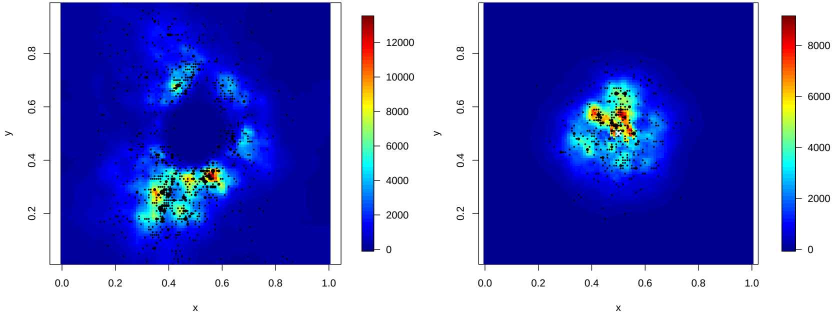
Example : Log-Gaussian Cox process

(stationary S process)



Example 2 : Log-Gaussian Cox processes

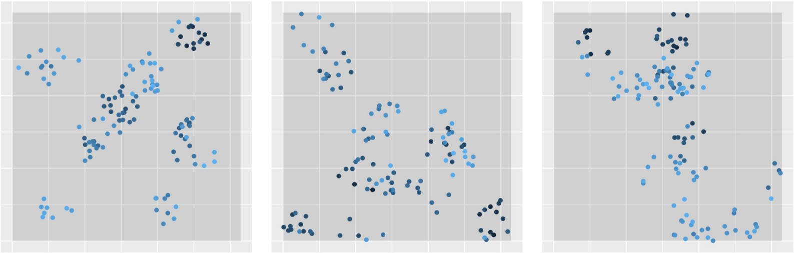
(nonstationary S process)



Example 3 : Neyman–Scott processes

taken from [González et al., 2016]

kernels : Gaussian, exponential, uniform (from left to right)
darker color indicates earlier t_i



Model classes with other construction principles

Determinantal processes [Lavancier et al., 2015] :

$$\lambda^{(k)}(y_1, \dots, y_k) = \det[C](y_1, \dots, y_k), \quad k = 1, 2, \dots$$

- $\det[C](y_1, \dots, y_k)$ covariance matrix determinant based on a ST covariance C
- pair correlation function = correlation of C, but proper treatment of ST yet open
- full likelihood inference possible yet requires some approximation tricks

Space-time birth-death processes [Renshaw, 2015] :

- typically formulated through conditional intensity $\lambda(s, t \mid X_{t_i < t})$
- can be pure birth, e.g. **self-exciting processes** (Hawkes processes) :

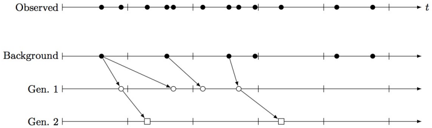


illustration from [Reinhart, 2017]

$\Rightarrow$ popular for earthquake modeling (ST) and in finance (T)

Two-step estimation

- use first and second-order information
- suitable when second moments (K, pair correlation) are easily available (Cox-like, determinantal, but not Gibbs)
- **Step 1 :** estimate $\lambda(s,t)$ (ST kernel smoothing, Poisson process model, ...)
- **Step 2 :** estimate interaction parameters
  - **minimum contrast :**

$$\hat{\theta} = \mathrm{argmin}_\theta \int_{D_{ST}} \left(K^c_{\text{emp}}(s,t) - K^c_\theta(s,t) \right)^2 \mathrm{d}(s,t)$$

    with calibration parameter $c > 0$
    $\Rightarrow$ nice theoretical results (consistency, asymptotic normality) [Waagepetersen and Guan, 2009]
  - **pairwise likelihood** with pairwise contributions

$$\ell(\theta; x_1, x_2) = \frac{\omega(x_1,x_2)\lambda^{(2)}(x_1,x_2)}{\int_{ST}\int_{ST}\omega(y_1,y_2)\lambda^{(2)}(y_1,y_2)\,\mathrm{d}y_1\mathrm{d}y_2}$$

    $\Rightarrow$ numerical integration, but very computer-intensive in ST
    see [Guan, 2006] for S processes

$$\text{Density}: \quad f(X) = C \exp\left(-\sum_{k=0}^{\infty} \sum_{i_1,\ldots,i_k} u_k(x_{i_1},\ldots,x_{i_k})\right), \quad C > 0$$

- **if orderly process** with known $\lambda^\star(s,t) = \lambda(s,t \mid X_{t_i<t}) \Rightarrow$ full likelihood

$$\exp\left(-\int_{D_{ST}} \lambda^\star(s,t)\,\mathrm{d}(s,t)\right) \prod_{i=1}^{N} \lambda^\star(s_i,t_i)$$

- otherwise, use Monte–Carlo likelihood to approximate C,
  or **pseudo-likelihood** with $\lambda^\star$ replaced by $\lambda^{\star\star}(x) = f(X \cup x)/f(X \setminus x)$.
- if $f(X) \propto \exp\left(-\sum_{\ell=1}^{m} \beta_\ell g_\ell(X)\right)$, discretize $\int_{D_{ST}} \ldots \Rightarrow$ **logistic regression** for β_ℓ
  [Baddeley and Turner, 2000, Baddeley et al., 2014], [Opitz, 2009] for 3D case
- **edge effects ?** i.e., interaction with unobserved points in D_{ST}^C
  $\Rightarrow$ **plus-sampling** on $D_{ST} \oplus r_0$ for simulation,
  minus-sampling on $D_{ST} \ominus r_0$ for estimation

Inference with log-Gaussian Cox processes

$W(s,t) = \log \Lambda(s,t)$ is a ST Gaussian random field

Likelihood is intractable :

$$\ell(W) = \mathbb{E} \exp\left(-\int_{D_{ST}} \exp(W(s,t))\, \mathrm{d}(s,t) \right) \prod_{i=1}^{N} \exp(W(s_i,t_i))$$

- **two-step estimation** possible : pair correlation function is $\exp(C(x_1,x_2))$
- **Bayesian inference :** hierarchical model for a Poisson process

$$\begin{aligned} X \mid \Lambda(s,t) &\sim \text{Poisson process}(\Lambda(s,t)) \\ \log \Lambda(s,t) &\sim \text{Gaussian process}(s,t;\mathrm{hyperpars}) \\ \mathrm{hyperpars} &\sim \ldots \end{aligned}$$

Hyperparameters determine variance and dependence in components of GP prior (variances in fixed and random effects, nugget, range, nonseparability, ...).

Bayesian inference with log-Gaussian Cox processes

$$\ell(W) = \mathbb{E} \exp\left(-\int_{D_{ST}} \exp(W(s,t))\,\mathrm{d}(s,t)\right) \prod_{i=1}^{N} \exp(W(s_i,t_i))$$

- $\mathbb{E}(\ldots)$ means integrating with respect to the density of the latent ST field $W(s,t)$
- **Markov-Chain Monte-Carlo :** iteratively simulate ST-discretized $W(s,t)$
  - use Gaussian approximation to posterior for Metropolis–Hastings proposals $\tilde{W}(s,t)$, see Section 4.4 of [Rue and Held, 2005]
    $\Rightarrow$ mixing may be slow
  - rather good mixing with MALA-MCMC : use gradient of log-posterior density [Taylor and Diggle, 2014]
- **Integrated Nested Laplace Approximation** [Rue et al., 2009, Opitz, 2017]
  - **Laplace approximation** for high-dimensional integrals
  - implemented in `R-INLA` for a large variety of generalized additive regression models
  - usually works fine in practice for log-Gaussian Cox processes
  - numerically efficient **Gauss–Markov models** for fast matrix calculations

MCMC and INLA provide ST posterior $\hat{\Lambda}(s,t)$ "for free" $\Rightarrow$ ST mapping

R packages for log-Gaussian Cox processes

`lgcp` [Taylor et al., 2011]

very complete functionality, including

- minimum-contrast estimation
- simulation-based estimation (MALA-MCMC) or interface to `RINLA`
- many models available (S models from `RandomFields` + other random effects), but limited for ST covariance

`RINLA` [Krainski et al., 2017, Rue et al., 2017]
see `http://www.r-inla.org/`

- available ST models :
  - temporal random effects : random walks, autoregression, ...
  - ST models : S innovations + temporal autoregression
    - S models : CAR, SPDE (= Gauss–Markov approximation to Matérn covariance)
    - some nonseparable SPDE models soon to come
  - packages providing additional features and easier syntax (but mainly S, not ST) :
    - `geostatsinla` [Brown, 2015]
    - `inlabru` : off-the-shelf fitting of log-Gaussian Cox processes for ecological applications

Recap of inference techniques

Frequentist :

- full likelihood for
  - Poisson models
  - orderly with known $\lambda(s, t \mid X_{t_i < t})$

  $\Rightarrow$ typically, ST discretization and/or numerical integration
- minimum contrast, pairwise likelihood
  $\Rightarrow$ for Cox-like models (known moments)
- pseudo-likelihood, Monte-Carlo likelihood
  $\Rightarrow$ for Gibbs-like models (unknown moments)

Bayesian :

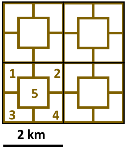
- simulation-based : MCMC
  $\Rightarrow$ for Cox-like models
- using analytical approximations : INLA
  $\Rightarrow$ for log-Gaussian Cox processes

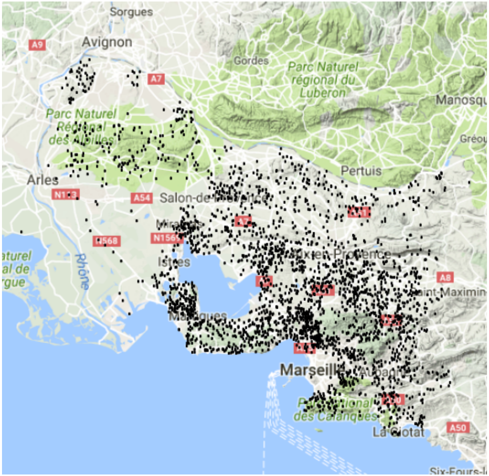
R implementations :

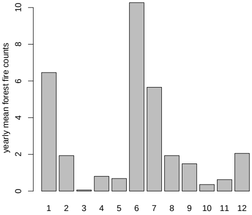
- `lgcp` and `RINLA` for log-Gaussian Cox processes
- `spatstat` for S processes
- `stpp` : focus on ST data and methodology, but parameter estimation yet to be developed

Wildfire reports from the Prométhée database

We focus on the **Bouches-du-Rhône department** :
$X = 2831$ reported wildfires (daily) from 1981-2015,
spatially indexed at the center of grid cells, with burnt area a_i.

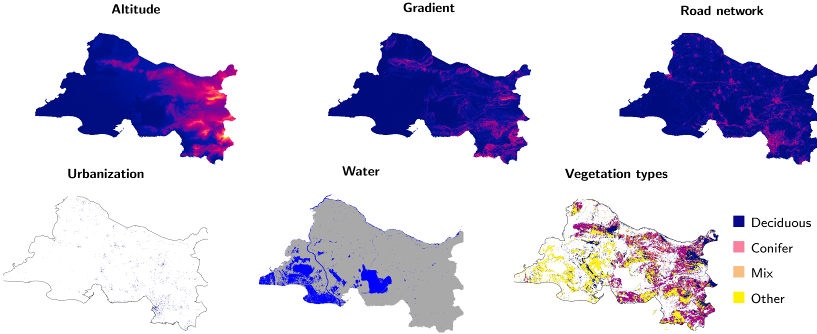






Relevant covariates

- **climate and weather :** temperature, precipitation, humidity, wind speed, ... (Salon-de-Provence weather station)
- **environmental** (IGN, 200m resolution grid)



Modeling approach

[Gabriel et al., 2017]

Goals :

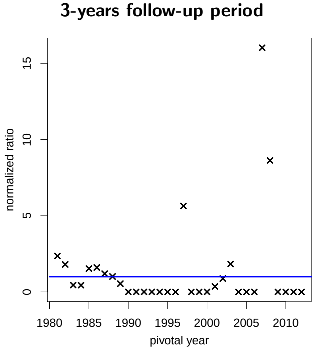
- **predict and map** wildfire risk over S and T
- detect characteristic **spatio-temporal structures**

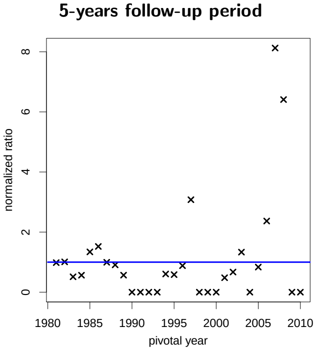
Modeling approach based on log-Gaussian Cox processes :

- **yearly** time step
- descriptive analyses of daily weather variables' impact $\Rightarrow$ yearly aggregation
- create synthetic covariate for **temporal inhibition at small spatial scales**
- check models for **residual space-time effects** (clustering, interaction)

Temporal inhibition at small spatial distance

- simple explanation (no combustible material left), but nice example with both clustering and repulsion at different scales
  $\Rightarrow$ ST extent of inhibitive effect ? (⚠ high positional uncertainty in s_i)
- **normalized empirical intensity ratio index** to confront expected/observed nearby event counts after an event $\Rightarrow$ value < 1 indicates repulsion





$\Rightarrow$ **construct a covariate** $z_{\text{inhib}}(s,t) = \sum_{i=1}^{N} (t_{\max} - (t - t_i)_+)_+ \times 1(\|s - s_i\| \leq r_{\max})$
with $t_{\max} = 5$ years and $r_{\max} = 1km$

Log-Gaussian Cox model

Stochastic intensity of log-Gaussian type

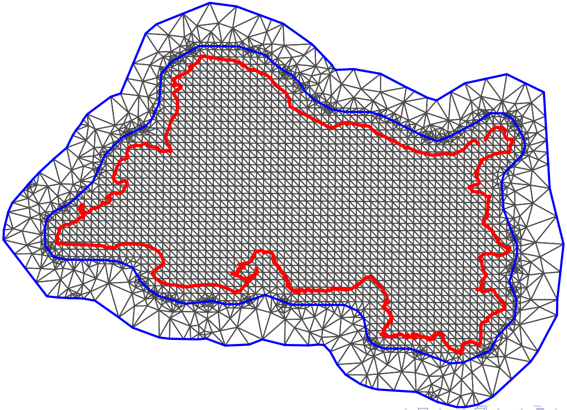
$$\Lambda(s,t) = \exp\left(\beta_0 + \sum_{c\in\texttt{clim}} \beta_c z_c(t) + \sum_{\ell\in\texttt{land use}} \beta_\ell z_\ell(s) + \beta_{\texttt{inhib}} z_{\texttt{inhib}}(s,t) + W(s,t)\right),$$

- **covariates :**
  - `clim` : related to proportion of days with exceedances of thresholds for mean temperature and deficits below thresholds for cumulated precipitation within the 4 weeks preceding the day
  - `land use` : proportion of land occupied by water, forest (conifer, deciduous), road and paths, buildings and average value of altitude and norm of the gradient
  - `inhib` : temporal inhibition effect (over 5 years)
- **Gaussian ST effect :**

$$\begin{aligned} W(s,t) &= W(s) && \text{S marginal effect,} && (1)\\ W(s,t) &= W(s) + W(t) && \text{S and T marginal effects,} && (2)\\ W(s,t) &= W_t(s) && \text{S effect i.i.d. in time for } t = 1981,\ldots,2015. && (3) \end{aligned}$$

- **inference with R-INLA** :
  - noninformative Gaussian priors on fixed effects β_i
  - S Matérn field $W(s)$ with smoothness $\nu = 1$

Time discretization = years 1981 to 2015
Triangulation of spatial domain to solve SPDE for Gauss–Markov approximation $W(s)$



Estimated covariate effects

$$\Lambda(s,t) = \exp\left(\beta_0 + \sum_{\texttt{clim}} \beta_c z_c(t) + \sum_{\texttt{land use}} \beta_\ell z_\ell(s) + \beta_{\texttt{inhib}} z_{\texttt{inhib}}(s,t) + W(s) + W(t)\right)$$

Covariate	**Posterior mean**	**95% credible interval**
intercept*	-7.3	[-8.06,-6.56]
inhib	-0.13	[-0.64,0.35]
clim1*	0.07	[0.05,0.1]
clim2*	0.57	[0.39,0.75]
clim3*	0.25	[0.2,0.31]
water*	-3.64	[-5.03,-2.36]
forest	0.59	[-0.01,1.19]
conif*	0.79	[0.19,1.4]
decid	-0.3	[-1.12,0.51]
alti*	-0.54	[-0.79,-0.28]
grad*	0.25	[0.11,0.38]
road*	0.49	[0.33,0.65]
path*	0.29	[0.12,0.47]
build	-0.04	[-0.1,0.01]
veget*	0.67	[0.13,1.22]

* *Significant effects.*

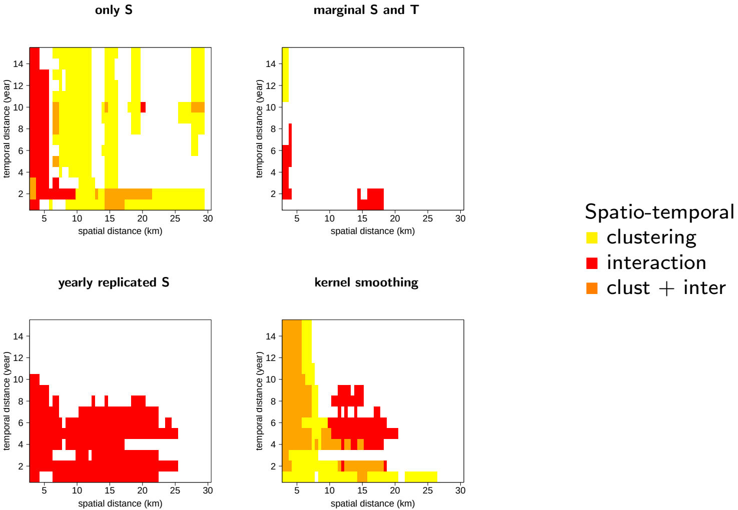
inhib not significant, but high positional uncertainty. Negative coefficient indicates tendency towards inhibition.

Hot and dry weather increases fire risk.

The presence of coniferous, dense road or path networks strongly favors wildfires
whereas
fire risk tends to significantly decrease with altitude and water.

Spatio-temporal variation and interaction

Fitted posterior intensity is separable in S and T.
We use the posterior mean for the log-Gaussian models.

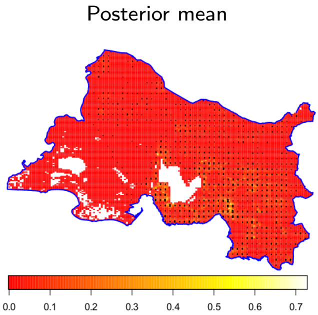


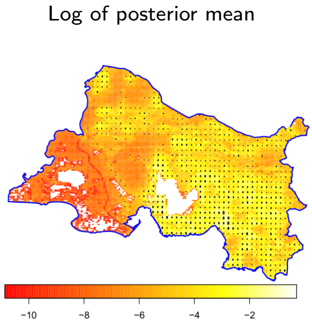
Taking into account the stochasticity in the posterior would further reduce zones with clustering/interaction.

Posterior means of spatial intensities of yearly wildfires (per km^2)

Purely S intensity model :

$$\Lambda(s,t) = \exp\left(\beta_0 + \sum_{\texttt{land use}} \beta_\ell z_\ell(s) + W(s)\right)$$

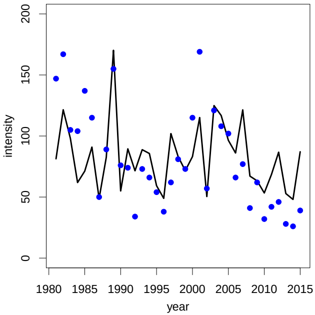




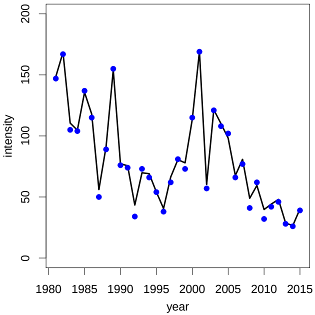
Estimated time trend

spatially averaged posterior means of temporal intensities of yearly wildfires

climate covariates only



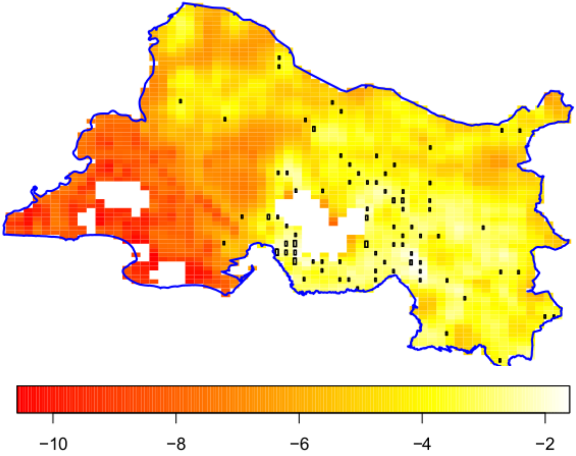
climate covariates and random walk



● observed number of wildfires for each year.

Prediction forward in time

Predicted spatial log-intensity for 2016 (per km^2)
based on model (2) with marginal S and T effects fitted with data until 2015
using climatic covariates of 2016 for prediction



• Wildfires observed in 2016 from January to October

Ongoing work for this application

- more complex ST dependence in $W(s,t)$ (auto-regression, nonseparable, ...)
- model validation/selection : DIC, WAIC, residual analysis, ...
- weekly, daily modeling and **seasonal effects**
- spatially varying climate covariates + **climate change scenarios**
- model the full PACA region, not only the Bouches-du-Rhône department
- consider **burnt areas as marks**

Some nice things not discussed here ...

- Gaussian-based Cox processes :
  if $\Lambda(s,t) = f(W(s,t)) \leq c < \infty$, can get rid of the ST discretization error
  $\Rightarrow$ MCMC mimicks random thinning of a Poisson process with rate c
  [Gonçalves and Gamerman, 2015]

- variational inference, EM inference

- **marked ST point processes** : mark m_i associated to point x_i
  - conditional to covariates,
    geostatistical marks are stochastically independent of the underlying points X
    $\Rightarrow$ separate modeling of X and $M = \{m_1, \dots, m_N\}$
  - BUT : marks may depend stochastically on the intensity
  - can consider the point process $\{(x_i, m_i)\}$ over $D_{ST} \times$ mark space
  - if marks are categorical $\Rightarrow$ multivariate point process

$\Rightarrow$ great number of possible models ...

Final remarks

- typical issues with complex ST data :
  - **heterogeneity** in spatial and temporal data types and observation scales
  - **curse of dimensionality :** $\dim(S) = 2$, $\dim(\mathrm{ST}) = 3$, $\dim(\mathrm{ST} + \text{marks}) > 3$
    - ⚠ nonseparable kernel smoothing
    - ⚠ inference often requires numerical integration
      e.g., Gibbs pseudo-likelihoods, ST discretization for log-Gaussian Cox processes
- in empirical modeling, **point counts may be aggregated** over larger ST units
  $\Rightarrow$ lack of information, but can use model structure similar to Cox processes
- many S models based on S **geometric objects** (kernels / interaction buffers)
  $\Rightarrow$ use ST objects, e.g. discs $\rightsquigarrow$ ST cylinders
  $\Rightarrow$ distinguish objects with **forward/backward/symmetric orientation in time**
- **model selection/validation tools :** visual checks on summary statistics, likelihood-based tools, residual analysis, cross-validation (easier in ST than in S)
- **a nice reference for ST point process inference :**

  González, J. A., Rodríguez-Cortés, F. J., Cronie, O., & Mateu, J. (2016). Spatio-temporal point process statistics : a review. *Spatial Statistics*, 18, 505-544.

Baddeley, A., Coeurjolly, J.-F., Rubak, E., and Waagepetersen, R. (2014).
Logistic regression for spatial Gibbs point processes.
Biometrika, 101(2) :377–392.

Baddeley, A. and Turner, R. (2000).
Practical maximum pseudolikelihood for spatial point patterns.
Australian & New Zealand Journal of Statistics, 42(3) :283–322.

Brown, P. E. (2015).
Model-based geostatistics the easy way.
Journal of Statistical Software, 63(12) :1–24.

Diggle, P. J. (2006).
Spatio-temporal point processes : methods and applications.

Gabriel, E., Opitz, T., and Bonneu, F. (2017).
Detecting and modeling multi-scale space-time structures : the case of wildfire occurrences.
Journal of the French Statistical Society (Special Issue on Space-Time Statistics).

Gonçalves, F. B. and Gamerman, D. (2015).
Exact Bayesian inference in spatiotemporal Cox processes driven by multivariate Gaussian processes.
Journal of the Royal Statistical Society : Series B (Statistical Methodology).

González, J. A., Rodríguez-Cortés, F. J., Cronie, O., and Mateu, J. (2016).
Spatio-temporal point process statistics : a review.
Spatial Statistics, 18 :505–544.

Guan, Y. (2006).
A composite likelihood approach in fitting spatial point process models.
Journal of the American Statistical Association, 101(476) :1502–1512.

Krainski, E. T., Lindgren, F., Simpson, D., and Rue, H. (2017).
The R-INLA tutorial on SPDE models Warning : work in progress... Suggestions to elias@r-inla. org are welcome.

Lavancier, F., Møller, J., and Rubak, E. (2015).
Determinantal point process models and statistical inference.
Journal of the Royal Statistical Society : Series B (Statistical Methodology), 77(4) :853–877.

Opitz, T. (2009).
Simulating and fitting of Gibbs processes in 3D – models, algorithms and their implementation.
Diploma Thesis, 121 pages.

Opitz, T. (2017).
Latent Gaussian modeling and INLA : A review with focus on space-time applications.
Journal of the French Statistical Society (Special Issue on Space-Time Statistics), 158(3).

Reinhart, A. (2017).
A review of self-exciting spatio-temporal point processes and their applications.
arXiv preprint arXiv :1708.02647.

Renshaw, E. (2015).
Stochastic population processes : analysis, approximations, simulations.
Oxford University Press.

Rue, H. and Held, L. (2005).
Gaussian Markov random fields : theory and applications.
CRC press.

Rue, H., Martino, S., and Chopin, N. (2009).

Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations.
Journal of the royal statistical society (Series B), 71(2) :319–392.

Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P., and Lindgren, F. K. (2017).
Bayesian computing with INLA : a review.
Annual Review of Statistics and Its Application, 4 :395–421.

Taylor, B. M., Davies, T. M., Rowlingson, B. S., and Diggle, P. J. (2011).
lgcp : An R Package for Inference with Spatio-Temporal Log-Gaussian Cox Processes.
arXiv preprint arXiv :1110.6054.

Taylor, B. M. and Diggle, P. J. (2014).
INLA or MCMC ? A tutorial and comparative evaluation for spatial prediction in log-Gaussian Cox processes.
Journal of Statistical Computation and Simulation, 84(10) :2266–2284.

Van Lieshout, M.-C. N. M. (2000).
Markov point processes and their applications.
World Scientific.

Van Lieshout, M.-C. N. M. (2006).
Markovianity in space and time.
Lecture Notes-Monograph Series, pages 154–168.

Waagepetersen, R. and Guan, Y. (2009).
Two-step estimation for inhomogeneous spatial point processes.
Journal of the Royal Statistical Society : Series B (Statistical Methodology), 71(3) :685–702.

Waagepetersen, R. P. (2007).
An estimating function approach to inference for inhomogeneous Neyman–Scott processes.
Biometrics, 63(1) :252–258.