Estimating the intensity function of spatial point processes outside the observation window

Edith Gabriel & co*

Laboratoire de Mathématiques d'Avignon

Rencontres de Statistique Avignon-Marseille
Marseille, June 2018

* J Chadœuf ; P Monestiez, F Bonneu ; F Rodriguez-Cortes, J Mateu ; J Coville

Motivations

Predicting the local intensity

Defining the predictor by a linear combination of the point process realization

Solving the Fredholm equation

to find the weights of the linear combination

$\Rightarrow$ approximated solutions

Work in progress

Discussion

About point processes

A point process, Φ on $\mathbb{R}^d$ is a random variable taking values in a measurable space $[\mathbb{X}, \mathcal{X}]$, where $\mathbb{X}$ is the family of all sequences φ of points of $\mathbb{R}^d$ satisfying

(i) the sequence is locally finite, i.e each bounded subset of $\mathbb{R}^d$ contains a finite number of points of φ.

(ii) the sequence is simple: $x_i \neq x_j$, if $i \neq j$.

Notations

- $\Phi_W = \Phi \cap W$: point process observed in $W \subset \mathbb{R}^2$
- $\Phi(B) = \sum_{x \in \Phi} \mathbb{I}_B(x)$: number of points of Φ within the set B.

Intensity function

Probability of one event within an elementary region:

$$\mathbb{P}[\text{ there is point of } \Phi \text{ in } dx\ ] = \lambda(x)\, \mathrm{d}x$$

where dx is an elementary region centered at x, with volume $\nu(dx)$.

$$\lambda(x) = \lim_{\nu(dx)\to 0} \frac{\mathbb{E}\left[\Phi(dx)\right]}{\nu(dx)}.$$

Inhomogeneity (i.e spatial variations) of the intensity can reflect:

- spatial variation in abundance (of a bird population), fertility (of a forest) or risk (of tornadoes),
- preference (of animal for certain types of habitat),
- dependence on external factors.

Pair correlation function

Relationship between number of events in a pair of subregions

$$g(x_i, x_j) = \frac{\lambda_2(x_i, x_j)}{\lambda(x_i)\lambda(x_j)}$$

where λ_2 is the second-order intensity function :

Probability of two events, each within an elementary region:

$$\mathbb{P}\left[\begin{array}{c}\text{one point of } \Phi \text{ in } dx_i \\ \text{and} \\ \text{one point of } \Phi \text{ in } dx_j\end{array}\right] = \lambda_2(x_i, x_j)\,\mathrm{d}x_i\,\mathrm{d}x_j$$

$$\lambda_2(x_i, x_j) = \lim_{\nu(dx_i)\to 0, \nu(dx_j)\to 0} \frac{\mathbb{E}\left[\Phi(dx_i)\Phi(dx_j)\right]}{\nu(dx_i)\nu(dx_j)}$$

Remark: for an isotropic process $\lim_{\|x_i - x_j\|\to\infty} g(x_i, x_j) = 1$,

since the events "there is a point of Φ in dx_i" and "there is a point of Φ in dx_j" are independent for large $\|x_i - x_j\|$.

Our aim

Let Φ a spatial point process, observed in a window W_{obs}.

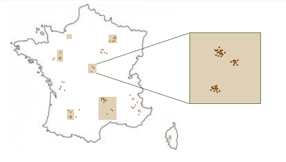
Can we predict its intensity function outside W_{obs}, conditionally to $\Phi \cap W_{obs}$?

Why? Exhaustive observations are impossible $\Rightarrow$ observation in quadrats.

Motivating example

How to estimate the spatial distribution of a bird species at a national scale from observations made in windows of few hectares?

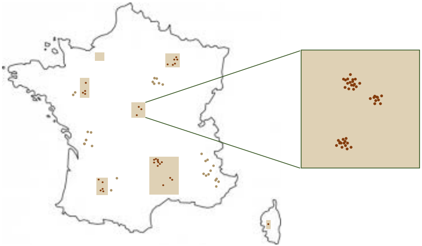
i.e. how to map *local intensity variations* of a point process in a large window when observation are available at a much smaller scale only?



Local intensity

Definition

We call *local* intensity of the point process Φ, its intensity conditional to its realization in W_{obs}: $\lambda(x|\Phi \cap W_{obs})$.



Window of interest:

$$\begin{aligned} W &= W_{obs} \cup W_{unobs} \\ &= (\cup \blacksquare) \cup (\cup \square) \end{aligned}$$

$$\Phi = \{\circ, \bullet\};\ \Phi_{W_{obs}} = \{\bullet\}$$

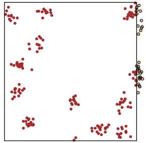
Our aim

To predict the local intensity in an unobserved window W_{unobs}.

Examples

Thomas process:

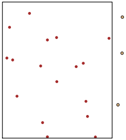
- κ: intensity of the Poisson process parents, Z,
- μ: mean number of offsprings per parent,
- σ: standard deviation of Gaussian displacement.



If W_{obs} splits a cluster, the local intensity across the boundary should be larger than λ.

Softcore process

If an event is observed close to the boundary of W_{obs}, the local intensity should be smaller the global one.



Existing solutions

- From the reconstruction of the process
  - Reconstruction method based on the 1^{st} and 2^{d}-order characteristics of Φ (see e.g. Tscheschel & Stoyan, 2006).
  - Get the intensity by kernel smoothing.

  A simulation-based method $\Rightarrow$ long computation times.

- For specific models
  - Diggle *et al.* (2007, 2013): Bayesian framework
  - Monestiez *et al.* (2006, 2013): Derived from classical geostatistics.

  Models constrained within the class of Cox processes.

- van Lieshout and Baddeley (2001).
  Based on exact simulations.

Our alternative approach

We want to predict the local intensity $\lambda(x|\Phi_{W_{obs}})$

- without precisely knowing the underlying point process model
  $\Rightarrow$ we only consider the 1^{st} and 2^{d}-order characteristics,
- in a reasonable time.

We define the predictor, similarly to a kriging interpolator, ie

- it is linear,
- it is unbiased,
- it minimizes the error prediction variance,

with weights depending on the structure of the point process.

Context

Let Φ a point process observed in W_{obs}.

For sake of clarity, we start by assuming that Φ is stationary[1], thus the global intensity and pair correlation function are

$$\lambda = \frac{\mathbb{E}\left[\Phi(W_{obs})\right]}{\nu(W_{obs})} \; ; \; g(x-y) = \frac{\lambda_2(x-y)}{\lambda^2}.$$

[1] *Assumption being relaxed later in the talk*

Our predictor

Proposition (Gabriel, Coville & Chadœuf, 2017)

For $x_o \in W_{obs}$,

$$\widehat{\lambda}(x_o|\Phi_{W_{obs}}) = \int_{\mathbb{R}^2} w(x; x_o) \sum_{y \in \Phi_{W_{obs}}} \delta(x-y)\,\mathrm{d}x = \sum_{x \in \Phi_{W_{obs}}} w(x; x_o)$$

is the Best Linear Unbiased Predictor of $\lambda(x_o|\Phi_{W_{obs}})$.
The weights, $w(x)$, are solution of the Fredholm equation of the 2^d kind:

$$w(x) + \lambda \int_{W_{obs}} w(y)\,(g(x-y)-1)\,\mathrm{d}y - \frac{1}{\nu(W_{obs})}\left[1 + \lambda \int_{W_{obs}^2} w(y)\,(g(x-y)-1)\,\mathrm{d}x\,\mathrm{d}y\right]$$
$$= \lambda\,(g(x_o - x) - 1) - \frac{\lambda}{\nu(W_{obs})} \int_{W_{obs}} (g(x_o - x) - 1)\,\mathrm{d}x$$

and satisfy $\int_{W_{obs}} w(x)\,\mathrm{d}x = 1$.
The variance of the predictor is given by

$$\mathbb{V}\mathrm{ar}\left(\widehat{\lambda}(x_o|\Phi_{W_{obs}})\right) = \lambda \int_{W_{obs}} w^2(x)\,\mathrm{d}x + \lambda^2 \int_{W_{obs} \times W_{obs}} w(x)w(y)\,(g(x-y)-1)\,\mathrm{d}x\,\mathrm{d}y.$$

Elements of proof

Linearity:

We set

$$\widehat{\lambda}(x_o|\Phi_{W_{obs}}) = \int_{\mathbb{R}^2} w(x;x_o) \sum_{y\in\Phi_{W_{obs}}} \delta(x-y)\,\mathrm{d}x = \sum_{x\in\Phi_{W_{obs}}} w(x;x_o).$$

Unbiasedness:

$$\mathbb{E}\left[\widehat{\lambda}(x_o|\Phi_{W_{obs}}) - \lambda(x_o|\Phi_{W_{obs}})\right] = 0$$

$$\begin{aligned}
&\iff \int_{W_{obs}} \lambda w(x)\,\mathrm{d}x - \mathbb{E}\left[\lim_{\nu(B)\to 0} \frac{\mathbb{E}\left[\Phi(B\oplus x_o)|\Phi_{W_{obs}}\right]}{\nu(B)}\right] = 0 \\
&\iff \lambda\left(\int_{W_{obs}} w(x)\,\mathrm{d}x - 1\right) = 0 \\
&\iff \int_{W_{obs}} w(x)\,\mathrm{d}x = 1.
\end{aligned}$$

Elements of proof

Minimum error prediction variance:

For any Borel set B,

$$\mathbb{V}\mathrm{ar}\left(\Phi(B)\right) = \lambda\nu(B) + \lambda^2 \int_{B\times B} \left(g(x-y)-1\right)\,\mathrm{d}x\,\mathrm{d}y$$

and for $B_o = B \oplus x_o$ with $x_o \notin W_{obs}$,

$$\lim_{\nu(B)\to 0} \frac{1}{\nu(B)} \int_{B_o \times W_{obs}} \left(g(x-y)-1\right)\,\mathrm{d}x\,\mathrm{d}y = \int_{W_{obs}} \left(g(x_o - x)-1\right)\,\mathrm{d}x$$

Then minimizing $\mathbb{V}\mathrm{ar}\left(\widehat{\lambda}(x_o|\Phi_{W_{obs}}) - \lambda(x_o|\Phi_{W_{obs}})\right)$ is equivalent to minimize

$$\lambda \int_{W_{obs}} w^2(x)\,\mathrm{d}x + \lambda^2 \int_{W_{obs}\times W_{obs}} w(x)w(y)\left(g(x-y)-1\right)\,\mathrm{d}x\,\mathrm{d}y \\ - 2\lambda^2 \int_{W_{obs}} w(x)\left(g(x_o-x)-1\right)\,\mathrm{d}x$$

Elements of proof

Using Lagrange multipliers under the constraint on the weights, we set

$$T(w(x)) = \lambda \int_{W_{obs}} w^2(x)\,\mathrm{d}x + \lambda^2 \int_{W_{obs}\times W_{obs}} w(x)w(y)\,(g(x-y)-1)\,\mathrm{d}x\,\mathrm{d}y$$
$$- 2\lambda^2 \int_{W_{obs}} w(x)\,(g(x_o - x) - 1)\,\mathrm{d}x + \mu\left(\int_{W_{obs}} w(x)\,\mathrm{d}x - 1\right)$$

Then, for $\alpha(x) = w(x) + \varepsilon(x)$,

$$T(\alpha(x)) \approx T(w(x)) + 2\lambda \int_{W_{obs}} \varepsilon(x)\left[w(x) + \lambda w(y)\,(g(x-y)-1)\,\mathrm{d}y\right.$$
$$\left. -\lambda\,(g(x_o - x) - 1) + \frac{\mu}{2\lambda}\right]\mathrm{d}x$$

Elements of proof

From variational calculation and the Riesz representation theorem,

$$\begin{aligned} T(\alpha(x)) \quad - \quad T(w(x)) &= o(\varepsilon(x)) \\ \Leftrightarrow \int_{W_{obs}} \varepsilon(x) \left[w(x)+\lambda\int_{W_{obs}} w(y)\,(g(x-y)-1)\,\mathrm{d}y - \lambda\,(g(x_o-x)-1) + \frac{\mu}{2\lambda}\right]\mathrm{d}x &= 0 \\ \Leftrightarrow \quad w(x) + \lambda \int_{W_{obs}} w(y)\,(g(x-y)-1)\,\mathrm{d}y - \lambda\,(g(x_o-x)-1) + \frac{\mu}{2\lambda} &= 0 \end{aligned}$$

Thus,

$$1 + \lambda \int_{W_{obs}^2} w(y)\,(g(x-y)-1)\,\mathrm{d}y\,\mathrm{d}x - \lambda \int_{W_{obs}} (g(x_o-x)-1)\,\mathrm{d}x + \frac{\nu(W_{obs})}{2\lambda}\mu = 0$$

from which we obtain μ and we can deduce the Fredholm equation

$$\begin{aligned} w(x)+\lambda \int_{W_{obs}} w(y)\,(g(x-y)-1)\,\mathrm{d}y - \frac{1}{\nu(W_{obs})}\left[1 + \lambda \int_{W_{obs}^2} w(y)\,(g(x-y)-1)\,\mathrm{d}x\,\mathrm{d}y\right] \\ = \lambda\,(g(x_o-x)-1) - \frac{\lambda}{\nu(W_{obs})} \int_{W_{obs}} (g(x_o-x)-1)\,\mathrm{d}x \end{aligned}$$

Solving the Fredholm equation

Any existing solution already considered in the literature can be used!

Our aim is to map the local intensity in a given window
$\Rightarrow$ access to fast solutions.

Several approximations can be used to solve the Fredholm equation.

The weights $w(x)$ can be defined as

- step functions $\rightsquigarrow$ direct solution,
- linear combination of known basis functions, e.g. finite elements, splines

  $\rightsquigarrow$ continuous approximation.
- ...

Here, we illustrate the ones with the less heavy calculations and implementation.

Finite element approach

The Fredholm equation can be written as

$$w(x) + \int_{W_{obs}} w(y)k(x,y)\,\mathrm{d}y = f(x;x_o), \tag{1}$$

with $k(x,y) = \lambda\left(g(x-y) - \frac{1}{\nu(W_{obs})}\int_{W_{obs}} g(x-y)\,\mathrm{d}x\right)$
and $f(x;x_o) = \frac{1}{\nu(W_{obs})} + \lambda\left(g(x-x_o) - \frac{1}{\nu(W_{obs})}\int_{W_{obs}} g(x-x_o)\,\mathrm{d}x\right)$.

The Galerkin method, with $\mathcal{T}_h$ a mesh partitioning W_{obs} and V_h an approximation space, plugged into a weak formulation of (1), leads to:

$$\sum_{j=1}^{N} w_j \int_{W_{obs}} \left(\varphi_i(x)\varphi_j(x) + \int_{W_{obs}}\int_{W_{obs}} k(x,y)\varphi_j(y)\varphi_i(x)\,\mathrm{d}y\right) = \int_{W_{obs}} f(x;x_o)\varphi_i(x)\,\mathrm{d}x,$$

with $w(x) \approx \sum_{i=1}^{N} w_i\varphi_i(x)$, $N = \dim V_h$ and $\{\varphi_i\}_{i=1,\dots,N}$ a basis of V_h.

Finite element approach

Using a matrix formulation, we have the Galerkin equation:

$$Mw + Kw = F, \tag{2}$$

with M the FEM mass matrix, $F = \left(\int_{W_{obs}} f(x; x_o) \varphi_i(x) \, \mathrm{d}x \right)_{i=1,\ldots,N}$

and $K = \left(\int_{W_{obs}} \int_{W_{obs}} k(x,y) \varphi_i(x) \varphi_j(y) \, \mathrm{d}x \, \mathrm{d}y \right)_{i,j}$.

We propose to solve (2) using $k(x,y) \approx \sum_{l,m} \mathcal{K}_{lm} \varphi_l(x) \varphi_m(y)$.

Thus, for $\mathcal{K} = (\mathcal{K}_{lm})_{l,m}$ and $K = M\mathcal{K}M$, this leads to consider the problem:

$$(Id + \mathcal{K}M)w = M^{-1}F, \tag{3}$$

When $\mathcal{T}_h$ is fine enough, (3) inherits the resolvability of the Fredholm equation, ensuring the consistency of the approximations.

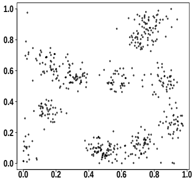
Finite element approach: illustrative results (1)

Simulation of a Thomas process within $[0,1] \times [0,1]$

Parents: $\mathcal{P}ois(\mu)$, $\mu = 50$

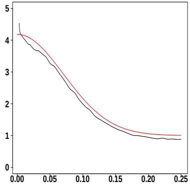
Offspring: $\mathcal{P}ois(\kappa)$, $\kappa = 10$, normally distributed, with $\sigma = 0.05$

Point process realization



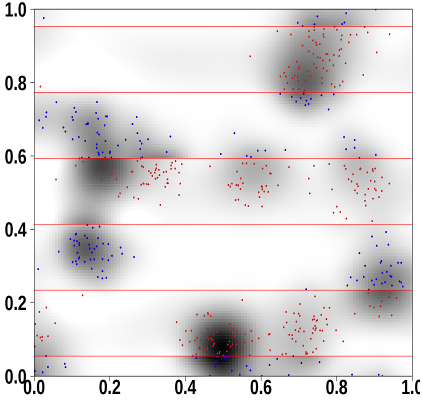
$\lambda = \kappa\mu = 500$

Pair correlation function



$g(r) = 1 + \frac{1}{4\pi\kappa\sigma^2} \exp\left(-\frac{r^2}{4\sigma^2}\right)$

Prediction within W_{unobs}

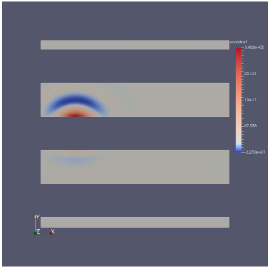


$\{\bullet\}$: $\Phi_{W_{obs}}$; $\{\bullet\}$: $\Phi_{W_{unobs}}$

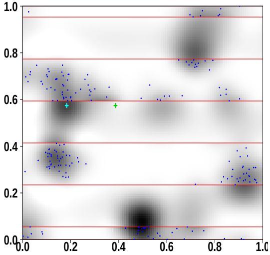
Finite element approach: illustrative results (1)

Weight function $w(\cdot; x_o)$

$x_o = (0.18, 0.57)$

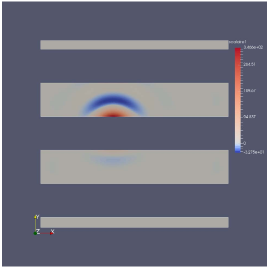


Prediction in W_{obs}



$\{\bullet\}$: $\Phi_{W_{obs}}$

$x_o = (0.38, 0.57)$



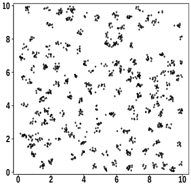
Finite element approach: illustrative results (2)

Simulation of a cluster process within $[0, 10] \times [0, 10]$

Parents: hardcore process with interaction radius 0.5

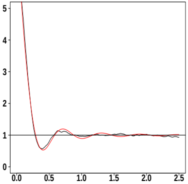
Offspring: normally distributed, with $\sigma = 0.1$

Point process realization



$\hat{\lambda} = 12.58$

Pair correlation function

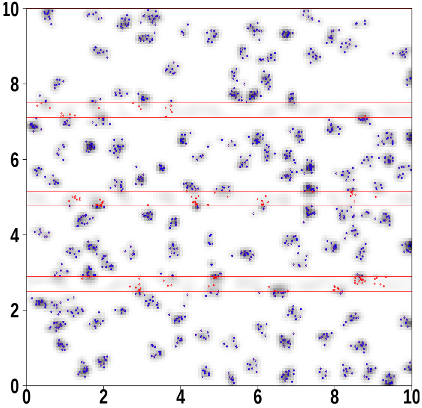


$$g(r) = 1 + \alpha \frac{\delta}{r} \exp\left(-\left(\frac{r}{\delta}\right)^{\beta}\right) \sin\left(\frac{r}{\delta}\right)$$

$$\hat{\alpha} = 11.65; \; \hat{\beta} = 0.35; \; \hat{\delta} = 1.25$$

$\{\bullet\}$: $\Phi_{W_{obs}}$; $\{\bullet\}$: $\Phi_{W_{unobs}}$

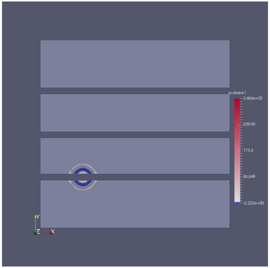
Prediction within W_{unobs}



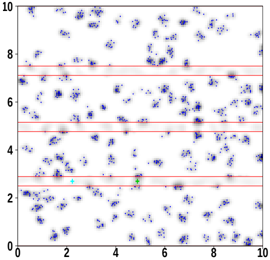
Finite element approach: illustrative results (2)

Weight function $w(\cdot; x_o)$

$x_o = (2.22, 2.69)$

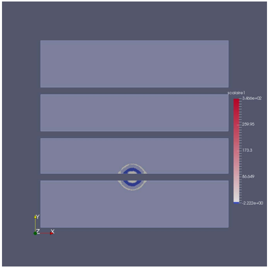


Prediction in W_{obs}



$\{\bullet\}$: $\Phi_{W_{obs}}$

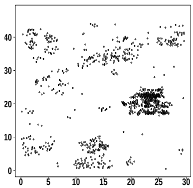
$x_o = (4.88, 2.69)$



Finite element approach: application

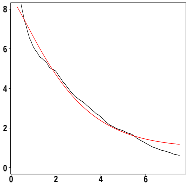
902 trees sampled in a 15*ha* quadrat to study the sahelian ecosystem[2].

Data location



$\hat{\lambda} = 0.6$

pcf

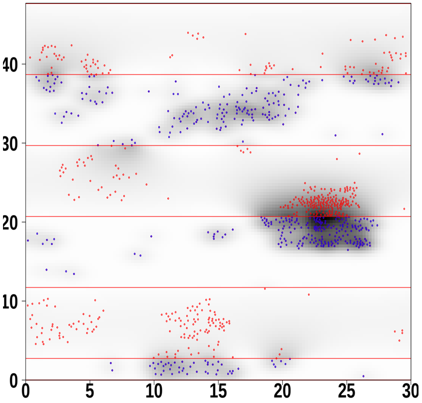


$g(r) = 1 + \alpha \exp\left(-\delta r^{\beta}\right)$

$\hat{\alpha} = 7.53$; $\hat{\beta} = 1.25$; $\hat{\delta} = 0.3$

$\{\bullet\}$: $\Phi_{W_{obs}}$; $\{\bullet\}$: $\Phi_{W_{unobs}}$

Prediction within W_{unobs}



[2] Dataset issued from an Intern. Biological Program conducted in Fété Olé, North Sénégal (Poupon, 1979)

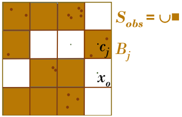
Step functions

Let us consider the following partition $S_{obs} = \cup_{j=1}^{n} B \oplus c_j$,

$B \oplus c_j$: elementary square centered at c_j,

$B \oplus c_k \cap B \oplus c_j = \emptyset$,

n: number of grid cell centers lying in S_{obs}.



Setting $w(x) = \sum_{j=1}^{n} w_j \frac{\mathbb{I}_{\{x \in B \oplus c_j\}}}{\nu(B)}$, leads to GBMC's predictor[3]:

$$\widehat{\lambda}(x_o|\Phi_{S_{obs}}) = \sum_{j=1}^{n} w_j \frac{\Phi(B \oplus c_j)}{\nu(B)}$$

with $w = (w_1, \dots, w_n) = C^{-1} C_o + \frac{1 - \mathbf{1}^T C^{-1} C_o}{\mathbf{1}^T C^{-1} \mathbf{1}} C^{-1} \mathbf{1}$, where

- $C = \lambda \nu(B) \mathbb{I} + \lambda^2 \nu^2(B)(G - 1)$: covariance matrix, with $\mathbb{I}$ the $n \times n$-identity matrix and $G = \{g_{ij}\}_{i,j=1,\dots,n}$, $g_{ij} = \frac{1}{\nu^2(B)} \int_{B \times B} g(c_i - c_j + u - v)\, \mathrm{d}u\, \mathrm{d}v$,
- $C_o = \lambda^2 \nu^2(B)(G_o - 1)$: covariance vector, with $G_o = \{g_{io}\}_{i=1,\dots,n}$.

[3] Gabriel, Bonneu, Monestiez & Chadœuf (2016)

Step functions: variance of the predictor

We consider the Neuman series to invert the covariance matrix, $C = \lambda\nu(B)\mathbb{I} + \lambda^2\nu^2(B)(G-1)$, when $\lambda\nu(B) \to 0$:

$$C^{-1} = \frac{1}{\lambda\nu(B)}\left[\mathbb{I} + \lambda\nu(B)J_\lambda\right],$$

where a generic element of the matrix J_λ is given by

$$J_\lambda[i,j] = \sum_{k=1}^{\infty}(-1)^k\lambda^{k-1}\left(g(x_i,x_{l_1})-1\right)\left(g(x_{l_{k-1}},x_j)-1\right)$$
$$\times\int_{W_{obs}^{k-1}}\prod_{m=1}^{k-2}\left(g(x_{l_m},x_{l_{m+1}})-1\right)\,\mathrm{d}x_{l_1}\ldots\,\mathrm{d}x_{l_{k-1}}.$$

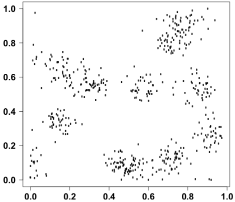
This leads to

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left(\widehat{\lambda}(x_o|\Phi_{W_{obs}})\right) &= \lambda^3\nu^2(B)(G_o-1)^T(G_o-1) + \lambda^4\nu^3(B)(G_o-1)^T J_\lambda(G_o-1)\\ &+\frac{1-\left[\lambda\nu(B)\mathbf{1}^T(G_o-1)+\lambda^2\nu^2(B)\mathbf{1}^T J_\lambda(G_o-1)\right]^2}{\frac{\nu(W_{obs})}{\lambda}+\nu^2(B)\mathbf{1}^T J_\lambda\mathbf{1}}.\end{aligned}$$

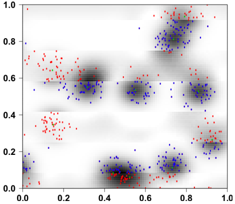
Step functions: illustrative results

Simulated Thomas process

$\kappa = 10, \mu = 50, \sigma = 0.05$



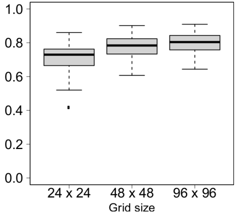
Prediction within W_{unobs}



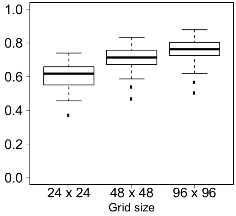
$\{\bullet\}$: $\Phi_{W_{obs}}$; $\{\bullet\}$: $\Phi_{W_{unobs}}$

R^2 in linear regression of predicted and theoretical values

(with the theoretical pcf)



(with the estimated pcf)



Two processes

Let $\Phi^{(1)}$ and $\Phi^{(2)}$ two stationary point processes observed in W_1 and W_2.

We want to predict the intensity of the Φ_1 at $x_o \notin W_1$ given $\Phi^{(1)}_{W_1}$ and $\Phi^{(2)}_{W_2}$.

We define

$$\widehat{\lambda_1}(x_o|\Phi^{(1)}_{W_1}, \Phi^{(2)}_{W_2}) = \sum_{x \in W_1} \omega_1(x) + \sum_{y \in W_2} \omega_2(y)$$

such that

$$\mathbb{E}\left[\widehat{\lambda}_1(x_o|\Phi^{(1)}_{W_1}, \Phi^{(2)}_{W_2})\right] = \lambda_1$$

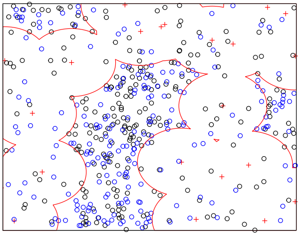
$$\mathbb{V}\mathrm{ar}\left(\widehat{\lambda}_1(x_o|\Phi^{(1)}_{W_1}, \Phi^{(2)}_{W_2}) - \lambda_1(x_o|\Phi^{(1)}_{W_1}, \Phi^{(2)}_{W_2})\right) \text{ minimum.}$$

$\rightsquigarrow$ a system of Fredholm equations.

$\rightsquigarrow$ depend on the cross pair correlation function.

Two processes: illustration

Multi-type Cox process driven by a boolean process of discs



- Generate a boolean process of discs

  Centers: from a Poisson process $\mathcal{P}(\lambda_b)$; Radius: R_b

- Generate two independent Poisson processes $\Phi^{(1)}_{init} \sim \mathcal{P}(\lambda_{o,1})$ and $\Phi^{(2)}_{init} \sim \mathcal{P}(\lambda_{o,2})$.

- Final processes: $\Phi^{(1)}$ and $\Phi^{(2)}$
  - Retain all points outside the union of discs,
  - Retain with probability p_i the points of $\Phi^{(i)}_{init}$ lying inside the union of discs.

Then, for $i,j \in 1,2 \quad \lambda_i = \lambda_{o,i}(e^{-\lambda_b \pi R_b^2} + (1-e^{-\lambda_b \pi R_b^2})p_i)$, and

$$g_{i,j}(r) = \frac{A + B(p_i+p_j) + (1-A-2B)p_i p_j}{(e^{-\lambda_b \pi R_b^2} + (1-e^{-\lambda_b \pi R_b^2})p_i)(e^{-\lambda_b \pi R_b^2} + (1-e^{-\lambda_b \pi R_b^2})p_j)}$$

with $A = e^{-\lambda_b S_r}$, $B = (1-e^{-\lambda_b(\pi R_b^2 - s_r)})e^{-\lambda_b \pi R_b^2}$,

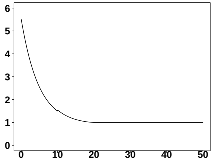
S_r (s_r): area of the union (intersection) of two discs of radii R_b, distant by r.

Two processes: illustration

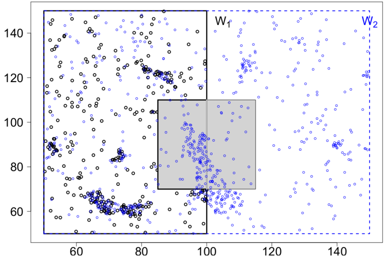
Parameters:

- Boolean process: $\lambda_b = 0.01$; $R_b = 10$,
- Poisson processes: $\lambda_{o,i} = 0.75$,
- Retention probabilities: $p_i = 0.05$.

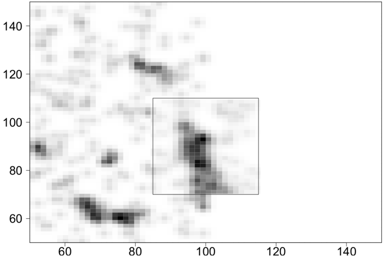
$g_{i,j}(r)$



Observations: $\Phi^{(1)}_{W_1}$ and $\Phi^{(2)}_{W_2}$



$\widehat{\lambda_1}(x_o|\Phi^{(1)}_{W_1}, \Phi^{(2)}_{W_2})$



Non-stationary processes

We relax the stationary assumption.

We assume that Φ is **S**econd-**O**rder **I**ntensity-**R**eweighted **S**tationary, i.e.

- its intensity $\lambda(x)$ is spatially varying
  e.g. it can be linked to covariates,

- the interaction between point depends on their difference (/distance):
$$g(x-y) = \frac{\lambda_2(x,y)}{\lambda(x)\lambda(y)}.$$

Non-stationary processes

The predictor has a similar definition: $\widehat{\lambda}(x_o|\Phi_{W_{obs}}) = \sum_{x \in \Phi_{W_{obs}}} w(x; x_o)$

The constraint on the weight function is

$$\int_{W_{obs}} \lambda(x) w(x) \, \mathrm{d}x = \lambda(x_o)$$

and the Fredholm equation becomes:

$$\begin{aligned} w(x) &+ \int_{W_{obs}} w(y)\lambda(y) \left(g(x-y)-1\right) \mathrm{d}y \\ &- \frac{1}{\nu(W_{obs})} \left[\int_{W_{obs}} w(x) \, \mathrm{d}x + \int_{W_{obs}^2} w(y)\lambda(y) \left(g(x-y)-1\right) \mathrm{d}x \, \mathrm{d}y \right] \\ &\qquad = \lambda(x_o) \left(g(x_o - x) - 1\right) - \frac{\lambda(x_o)}{\nu(W_{obs})} \int_{W_{obs}} \left(g(x_o - x) - 1\right) \mathrm{d}x \end{aligned}$$

Non-stationary processes: goodness of prediction

Let Φ be a SOIRS Neyman-Scott process obtained by $p(x)$-thinning, with

- intensity $\lambda(x) = \kappa\mu p(x)$,
- $\Phi^{(p)} \sim \mathcal{P}(\kappa)$ the process of parents,
- $f(x; R)$ the dispersion kernel for the offspring, with range R,
- mean number of offspring μ.

For $\partial W = W_{\oplus r} \backslash W$, the local intensity is

$$\lambda(x_o|\Phi_W) = \int \left[\sum_{y \in b(x_o,R)\cap(W\cup\partial W)} \mu p(x_o) f(y - x_o) + \mu\kappa \int_{b(x_o,R)\backslash(W\cup\partial W)} p(x_o) f(y - x_o)\, \mathrm{d}y \right] \mathrm{d}P[\Phi^{(p)}_{W\cup\partial W}|\Phi_W]$$

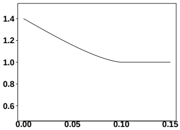
Then, we can get a Monte Carlo approximation of $\lambda(x_o|\Phi_W)$ by simulating K realizations of parent points in ∂W.

Non-stationary processes: illustration

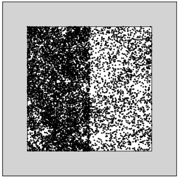
Independent thinning of a Matérn cluster process:

- Thinning probability:
  $p(x) = p(x_1, x_2) = 0.8\, \mathbb{I}_{\{x_1 \leq 0.5\}} + 0.2\, \mathbb{I}_{\{x_1 > 0.5\}}$.
- Offspring x are uniformly distributed on a disc of radius R around its parent point x_p :
  $f(x) = \frac{1}{\pi R^2} \mathbb{I}_{\{\|x - x_p\| \leq R\}}$; $R = 0.05$.
- $\kappa = 1000$, $\mu = 20$.

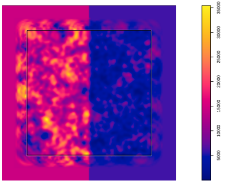
$g(r)$



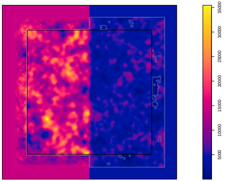
Observations: $\Phi_{W_{obs}}$



Prediction: $\widehat{\lambda}(x_o | \Phi_{W_{obs}})$



Conditional intensity



Forthcoming work: extend to spatio-temporal processes

For $(x_o, t_o) \notin D_{obs} = W_{obs} \times T_{obs}$, the spatio-temporal predictor given by $\hat{\lambda}((x_o, t_o)|\Phi_{D_{obs}}) = \sum_{(x,t)\in\Phi\cap D_{obs}} w(x,t)$ is the BLUP of $\lambda((x_o, t_o)|\Phi_{D_{obs}})$.

Assuming Φ stationary, $w(x,t)$ satisfies $\int_{D_{obs}} w(x,t)\,\mathrm{d}x\,\mathrm{d}t = 1$, and is solution of the Fredholm equation of the second kind:

$$\lambda(g(x_o - x_1, t_o - t_1) - 1) - \frac{\lambda}{\nu(D_{obs})} \int_{D_{obs}} w(x_1, t_1)(g(x_o - x_1, t_o - t_1) - 1)\,\mathrm{d}x_1\,\mathrm{d}t_1$$

$$= w(x_1, t_1) + \lambda \int_{D_{obs}} w(x_2, t_2)(g(x_1 - x_2, t_1 - t_2) - 1)\,\mathrm{d}x_2\,\mathrm{d}t_2$$

$$- \frac{1}{\nu(D_{obs})}\left[1 + \int_{D_{obs}\times D_{obs}} w(x_2, t_2)(g(x_1 - x_2, t_1 - t_2) - 1)\,\mathrm{d}(x_1, x_2)\,\mathrm{d}(t_1, t_2)\right].$$

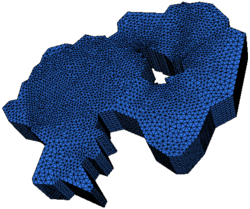
Forthcoming work: extend to spatio-temporal processes

Solve the Fredholm equation using the finite element approach:

$$w(x,t) \approx \sum w_i \varphi_i(x,t),$$

(should work because $D_{obs} = W_{obs} \times T_{obs}$).

Extend to SOIRS [4] spatio-temporal processes.



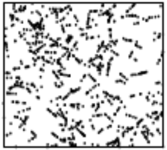
[4] Space-time Second-Order Intensity Reweighted Stationarity, see Gabriel & Diggle (2009)

Forthcoming work: extending to fibre processes

Applying the same approach to fibre processes:

- $\Rightarrow$ switch from summation on points to integral along fibres
- Again with the pair correlation function
  $\rightarrow$ local fibre orientation weakly taken into account.

The problem can also occur for point processes, e.g. for a Cox process driven by a boolean segment process



Forthcoming work: predicting the local intensity from data of different kind

Consider both

- data point locations, $\Phi_{W_1} = \Phi \cap W_1$,
- count data, $\Phi(W_2)$.

$\Rightarrow$ How to predict $\lambda(x_o|\Phi_{W_1}, \Phi(W_2))$, $x_o \notin W_1 \cup W_2$?

A (very) first candidate:

$$\widehat{\lambda}(x_o|\Phi_{W_1}, \Phi(W_2)) = \sum_{x \in \Phi \cap W_1} w(x) + \alpha \Phi(W_2).$$

to be continued ...

References

E. Bellier *et al.* (2013) Reducting the uncertainty of wildlife population abundance: model-based versus design-based estimates. *Environmetrics*, 24(7):476–488.

P. Diggle *et al.* (2013) Spatial and spatio-temporal log-gaussian cox processes: extending the geostatistical paradigm. Statistical Science, 28(4):542–563.

E. Gabriel, J. Coville, J. Chadœuf (2017) Estimating the intensity function of spatial point processes outside the observation window. *Spatial Statistics*, 22(2), 225–239.

E. Gabriel, F. Bonneu, P. Monestiez, J. Chadœuf (2016) Adapted kriging to predict the intensity of partially observed point process data. *Spatial Statistics*, 18, 54–71.

E. Gabriel, P. Diggle (2009) Second-order analysis of inhomogeneous spatio-temporal point process data. *Statistica Neerlandica*, 63, 43–51.

P. Monestiez *et al.* (2006) Geostatistical modelling of spatial distribution of balaenoptera physalus in the northwestern mediterranean sea from sparse count data and heterogeneous observation efforts, *Ecological Modelling*, 193:615–628.

A. Tscheschel and D. Stoyan (2006) Statistical reconstruction of random point patterns. *Computational Statistics and Data Analysis*, 51:859–871.

M-C. van Lieshout and A. Baddeley (2001) Extrapolating and interpolating spatial patterns, *In Spatial Cluster Modelling*, A.B. Lawson and D.G.T. Denison (EDS.) Boca Raton: Chapman And Hall/CRC, pp 61–86.