Statistical analysis of point patterns on linear networks

Jesper Møller, Aalborg University

This talk is based on

Møller, J. and J.G. Rasmussen (2022). Cox processes driven by transformed Gaussian processes on linear networks.
Scandinavian Journal of Statistics
(invited contribution; in preparation).

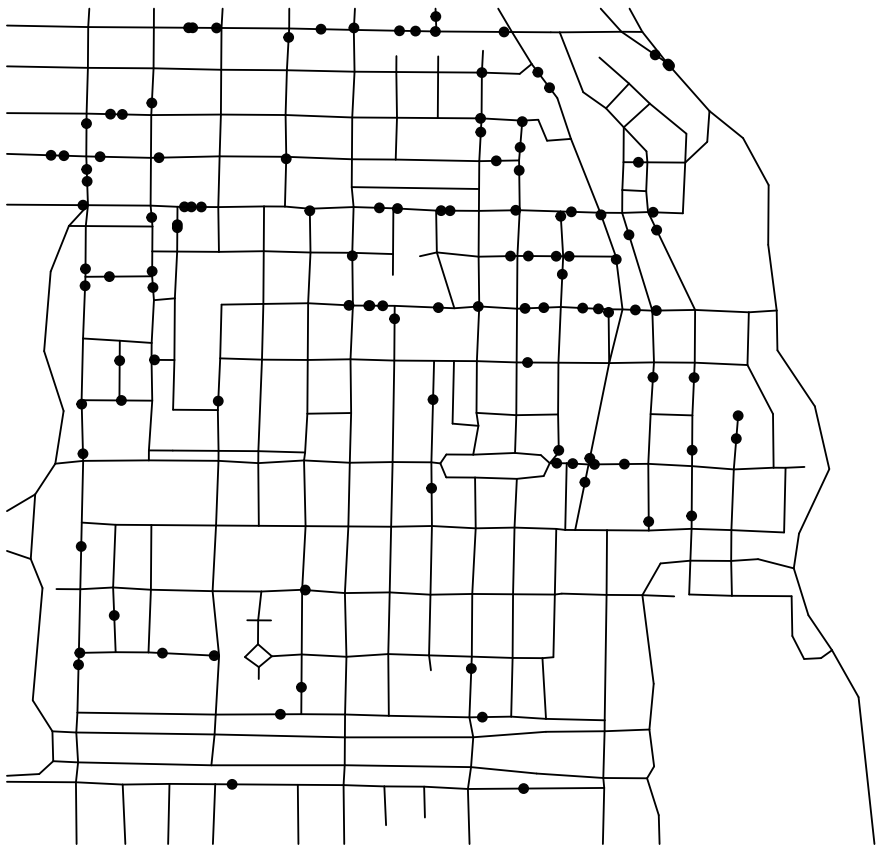
Available at https://people.math.aau.dk/~jm/courses/SummerSchool2022/sjs.pdf

We used the spatstat R-package together with our own software available at ....

The slides are available at https://people.math.aau.dk/~jm/courses/SummerSchool2022/lecture_networks_no_pause.pdf

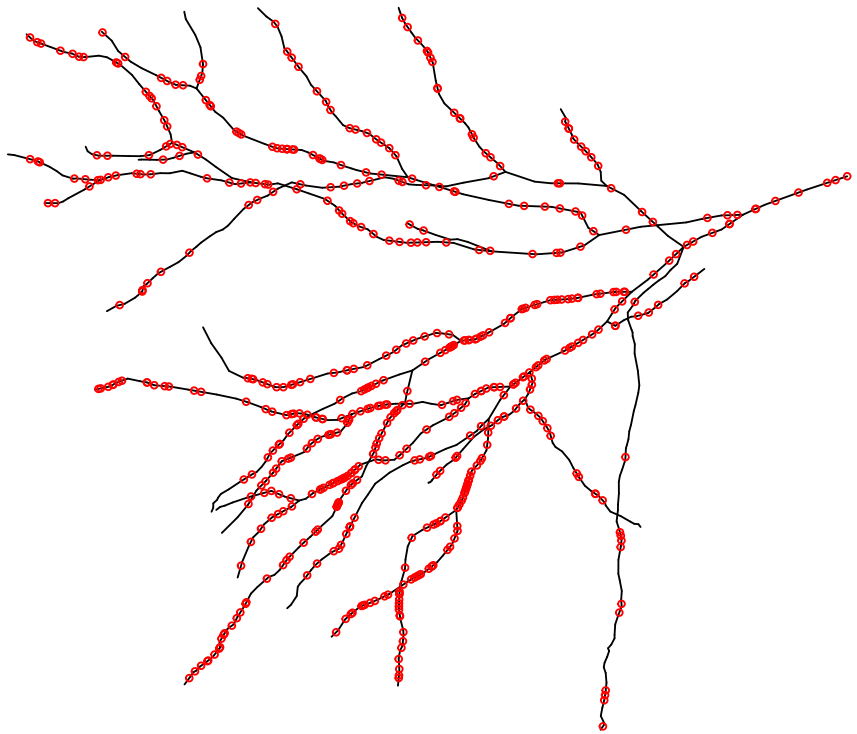
Data examples

Crimes in a part of Chicago



566 spines observed on one branch of the dendritic tree of a rat neuron

Dendrites are branching filaments which extend from the main body of a neuron (nerve cell) to propagate electrochemical signals. Spines are small protrusions on the dendrites.



Linear network

$L = \cup_{i=1}^{m} L_i$ is a linear network if

- $m < \infty$,
- each $L_i \subset \mathbb{R}^k$ is a closed line segment of length $l_i \in (0, \infty)$,
- $L_i \cap L_j$ is either empty or an endpoint of both L_i and L_j whenever $i \neq j$,
- L is a path-connected set.

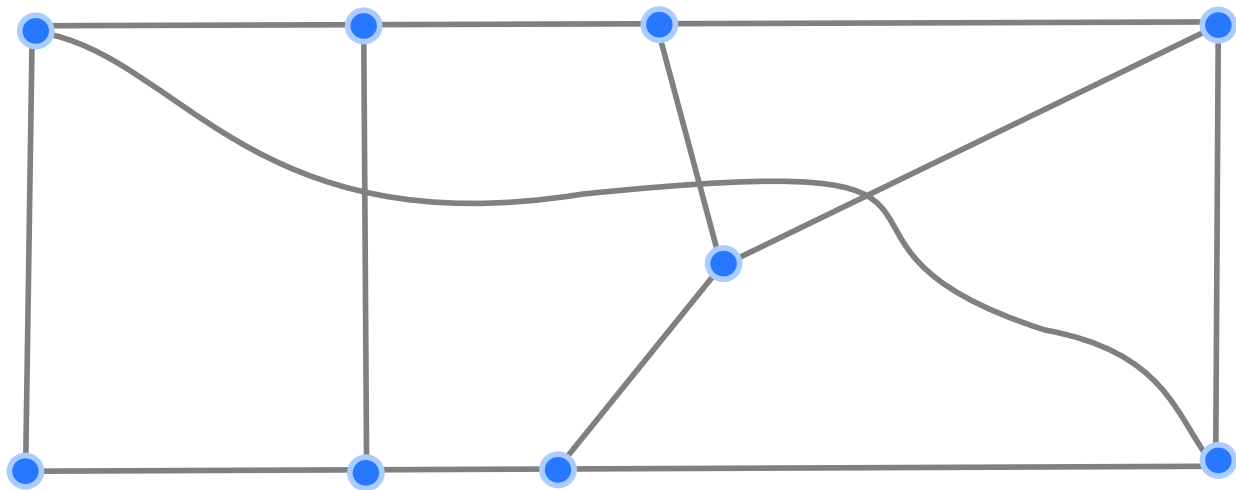
Natural measure is arc length measure d_L: For (Borel) sets $A \subseteq L$,

$$\text{arc-length}(A) = \int_A \mathrm{d}_L(u) = \sum_{i=1}^{m} \text{length}(A \cap l_i).$$

Extension

This may be extended to the more abstract case of a graph with Euclidean edges (Anderes, Møller & Rasmussen, 2020), i.e., graphs with edges viewed as curve segments and allowing tunnels and bridges...

A graph with Euclidean edges (not a linear network):



We avoid this generalization for ease of presentation and since statistical methods have (so far) only been developed for the case of linear networks.

Briefly about point process models (more later)

There has been a particular focus on developing point process models on L where the pair correlation function is isotropic with respect a given metric d on L: $g(x, y) = g_0(d(u, v))$, $u, v \in L$.

This is satisfied for a Poisson process ($g = 1$) but not for other Gibbs point processes because of boundary effects.

The use of $d(u, v) = \|u - v\|$ (Euclidean metric) is usually unnatural.

Though it is often sensible to $d =$ geodesic metric (Okabe & Sugihara, 2012), then 'point process models with an isotropic pcf are rare' (Baddeley, Nair, Rakshit & McSwiggan, 2017).

Later in this talk: construct various Cox process models s.t. g becomes isotropic but d needs to be something else ...

Metrics

We have two cases of natural metrics in mind: d is

- the geodesic metric $d_{\mathcal{G}}$
- or the resistance metric $d_{\mathcal{R}}$ (Anderes, Møller & Rasmussen, 2020): In brief, viewing L as an electrical network with resistor l_i at edge L_i, $i = 1, \ldots, m$, then $d_{\mathcal{R}}(u, v)$ is the effective resistance between u and v as obtained by Kirkhoff's laws. (???? Technical details follow !!!!)

As we shall see, $d_{\mathcal{R}}$ has many advantages over $d_{\mathcal{G}}$ when

- developing models for isotropic covariance functions/Gaussian processes
- used for a number of point process models with g isotropic w.r.t. $d_{\mathcal{R}}$ (and also $d_{\mathcal{G}}$ if L is a tree).

WARNING

The following 6 pages of slides provide the technical details on how to define and calculate $d_{\mathcal{R}}$.

This is just to make the technical inclined audience happy – and actually also to show that it is not difficult to make calculations.

Do not worry if you get lost – then just accept there is a definition and a way of making calculations.

I'll not spend time on it today!

Classic resistance metric $d_{\mathcal{V}}$ on the vertex set

The linear network induces a graph $G = (V, E)$.

The resistance metric on L ($d_{\mathcal{R}}$) is an extension of the classic (effective) resistance metric $d_{\mathcal{V}}$ on V: For every $u, v \in V$,

$$d_{\mathcal{V}}(u, v) = \mathbb{V}\mathrm{ar}(Z_0(u) - Z_0(v)) = \Sigma(u, u) + \Sigma(v, v) - 2\Sigma(u, v)$$

where $Z_0 = \{Z_0(u) \,|\, u \in V\}$ is Gaussian with mean zero and covariance matrix $\Sigma = \Delta^{-1}$ with Δ a 'Laplacian matrix' defined as follows.

Classic resistance metric $d_{\mathcal{V}}$ (cont.)

Let $\sim$ be the neighbour relation on V and $u_0 \in V$ an arbitrarily chosen vertex (the 'origin') and define a 'conductance function' by

$$\mathrm{con}(u, v) = 1/\|u - v\| \text{ if } u \sim v, \quad \mathrm{con}(u, v) = 0 \text{ otherwise,}$$

and the sum of the conductances associated to the edges incident to vertex u,

$$c(u) = \sum_{w \in V:\, w \sim u} \mathrm{con}(u, w),$$

and Δ by

$$\begin{aligned} \Delta(u_0, u_0) &= 1 + c(u_0), \\ \Delta(u, u) &= c(u) && \text{if } u \neq u_0, \\ \Delta(u, v) &= -\mathrm{con}(u, v) && \text{if } u \neq v. \end{aligned}$$

Resistance metric $d_{\mathcal{R}}$ on L

Extend Z_0 by linear interpolation to a mean zero Gaussian process (GP) on L: If the edge L_i has endpoints a_i and b_i,

$$Z_0(u) = \frac{\|u - b_i\|}{l_i} Z_0(a_i) + \frac{\|u - a_i\|}{l_i} Z_0(b_i) \quad \text{for } u \in L_i.$$

Let $Z_1, \ldots, Z_m$ be independent mean zero GPs so that $Z_i = 0$ outside L_i, and Z_i on L_i is a Brownian bridge with

$$\mathbb{C}\mathrm{ov}(Z_i(u), Z_i(v)) = \min\{\|u-a_i\|\|v-b_i\|, \|v-a_i\|\|u-b_i\|\}/l_i.$$

Then, for every $u, v \in L$,

$$d_{\mathcal{R}}(u, v) = \mathbb{V}\mathrm{ar}\left(\sum_{i=0}^{m} Z_i(u) - Z_i(v)\right).$$

Calculation of $d_{\mathcal{R}}$

For any $u \in L_j$ and $v \in L_i$, let

$$s = \|u - a_j\|, \quad t = \|v - a_i\|, \quad A_i = d_{\mathcal{V}}(a_i, b_i)/l_i^2 - 1/l_i \leq 0$$

where $A_i = 0 \Leftrightarrow L_i$ is the only path connecting a_i and b_i.

(A) If $i = j$ then

$$\begin{aligned} d_{\mathcal{R}}(u, v) = 1(t \geq s) \left[A_i(t - s)^2 + t - s\right] \\ + 1(t \leq s) \left[A_i(t - s)^2 + s - t\right]. \end{aligned}$$

As a function of t: linear ($A_i = 0$) or quadratic ($A_i < 0$) on $[0, s]$ and $[s, l_i]$, continuous, differentiable except for $t = s$.

Calculation of $d_{\mathcal{R}}$ (cont.)

(B) If $i \neq j$ then

$$d_{\mathcal{R}}(u, v) = A_i t^2 + B_{ij}(s)t + C_{ij}(s)$$

where

$$B_{ij}(s) = 1 - \frac{2}{l_i}[\Sigma(a_i, a_i) - \Sigma(a_i, b_i) - \frac{l_j - s}{l_j}\Sigma(a_j, a_i) + \frac{l_j - s}{l_j}\Sigma(a_j, b_i) - \frac{s}{l_j}\Sigma(b_j, a_i) + \frac{s}{l_j}\Sigma(b_j, b_i)]$$

and

$$C_{ij}(s) = \frac{(l_j - s)^2}{l_j^2}\Sigma(a_j, a_j) + \frac{s^2}{l_j^2}\Sigma(b_j, b_j) + 2\frac{s(l_j - s)}{l_j^2}\Sigma(a_j, b_j) + \Sigma(a_i, a_i) - 2\frac{l_j - s}{l_j}\Sigma(a_j, a_i) - 2\frac{s}{l_j}\Sigma(b_j, a_i) + \frac{s(l_j - s)}{l_j}.$$

Some results

If all vertices are of order two, L is isomorphic to a circle, so we say that L is a loop.

If there is no loop, we say that L is a tree.

Theorem

(A) *The definition of $d_{\mathcal{V}}$ does not on the choice of origin $u_0 \in V$, and $d_{\mathcal{R}}(u, v) = d_{\mathcal{V}}(u, v)$ if $u, v \in V$.*

(B) *Both $d_{\mathcal{R}}$ and $d_{\mathcal{G}}$ are metrics on L, and their definitions are invariant to splitting a line segment L_i into two line segments.*

(C) *$d_{\mathcal{G}} \geq d_{\mathcal{R}}$, with equality if and only if L is a tree.*

(D) *If G is a loop, then $d_{\mathcal{R}}(u, v) = d_{\mathcal{G}}(u, v) - d_{\mathcal{G}}(u, v)^2 / \sum_{i=1}^{m} l_i$.*

Gaussian processes on L

Let $Y = \{Y(u) \,|\, u \in L\}$ be a (real) Gaussian process (GP).

It is determined by the mean function $\mu(u) = \mathbb{E}Y(u)$ and the covariance function

$$c(u, v) = \mathbb{C}\mathrm{ov}(Y(u), Y(v)).$$

It is well-defined iff c is symmetric and positive definite, i.e.,

$$\sum_{j,\ell=1}^{n} a_j a_\ell c(u_j, u_\ell) \geq 0 \quad \text{for all } a_1, \ldots, a_n \in \mathbb{R},\ n = 1, 2, \ldots$$

Isotropic covariance functions

Want

$$c(u, v) = c_0(d(u, v)) \quad \text{for all } u, v \in L$$

since then we can construct isotropic GPs and hence – as we shall see – point process models with isotropic pair correlation functions.

We assume $\sigma^2 = c_0(0) > 0$ and consider the correlation function

$$r_0(t) = c_0(t)/\sigma^2$$

meaning that $r_0(d(u, v))$ has to be strictly positive definite over $(u, v) \in L \times L$.

Examples of parametric models

$r_0(d_{\mathcal{R}}(u,v))$ is strictly positive definite over $(u,v) \in L \times L$ in the following cases (conditions on L are needed if instead distance is measured by $d_{\mathcal{G}}$):

Model	Correlation function $r_0(t)$	Range of parameters
Powered exponential	$\exp\left(-t^{\alpha}/\phi\right)$	$0 < \alpha \leq 1$
Matérn	$\frac{2^{1-\alpha}}{\Gamma(\alpha)}\left(\sqrt{2\alpha}\frac{t}{\phi}\right)^{\alpha} K_{\alpha}\left(\sqrt{2\alpha}\frac{t}{\phi}\right)$	$0 < \alpha \leq \frac{1}{2}$
Generalized Cauchy	$(1+(\frac{t}{\phi})^{\alpha})^{-\tau/\alpha}$	$\tau > 0$, $0 < \alpha \leq 1$
Dagum	$1-((\frac{t}{\phi})^{\tau}/(1+(\frac{t}{\phi})^{\tau}))^{\frac{\alpha}{\tau}}$	$0 < \tau \leq 1$, $0 < \alpha \leq 1$

Table 1: Here, Γ is the gamma function, K_{ν} is the modified Bessel function of the second kind, $\phi > 0$ is a scale parameter, τ is a shape parameter, α is a smoothness parameter, and the range of (τ, α) depends on the model.

Isotropic and completely monotone covariance functions

In all examples, r_0 will be a completely monotone function, i.e.,

$$r_0 \geq 0 \text{ and is continuous on } [0, \infty)$$

and for $j = 1, 2, \ldots$ and all $u > 0$

$$\text{the } j\text{-th derivative } r_0^{(j)}(u) \text{ exists and } (-1)^j r_0^{(j)}(u) \geq 0.$$

Bernstein's theorem: r_0 is completely monotone if and only if it is the Laplace transform of a non-negative random variable S: for every $t \geq 0$,

$$r_0(t) = \mathbb{E} \exp(-tS) \quad \text{where } S \geq 0.$$

Examples...

... of completely monotone correlation functions r_0 and corresponding S:

For 1)-2) $\tau > 0$, $\phi > 0$ and 3) $\psi > 0$, $\chi > 0$, $\lambda \in \mathbb{R}$:

1) $r_0(t) = (1 + t/\phi)^{-\tau}, \quad S \sim \Gamma(\tau, \phi)$ (inverse scale parameter ϕ).

2) $r_0(t) = \dfrac{2\phi^\tau}{\Gamma(\tau)}(t/\phi)^{\tau/2} K_\tau(2\sqrt{t\phi}), \quad S \sim \Gamma^{-1}(\tau, \phi).$

3) $r_0(t) = (1 + 2t/\psi)^{-\lambda/2} \dfrac{K_\lambda(\sqrt{(2t+\psi)\chi})}{K_\lambda(\sqrt{\psi\chi})}$

and $S \sim$ generalized inverse Gaussian distribution with pdf

$$\frac{(\psi/\chi)^{\lambda/2}}{2K_\lambda(\sqrt{\psi\chi})} s^{\lambda-1} \exp(-s\psi/2 - \chi/(2s)), \quad s \geq 0.$$

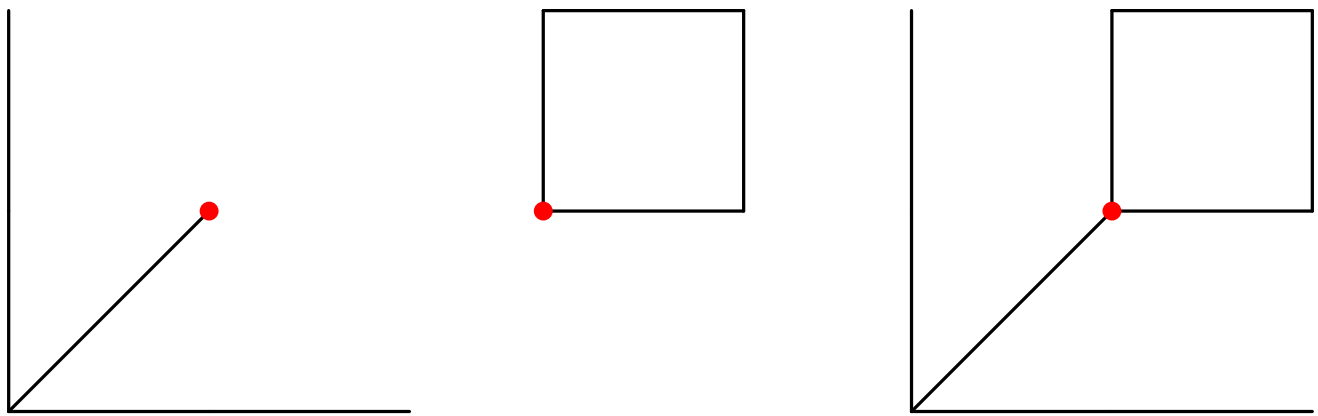
1-sums of linear networks

L is a 1-sum of $\mathcal{L}_1 = L_1 \cup \ldots \cup L_j$ and $\mathcal{L}_2 = L_{j+1} \cup \ldots \cup L_m$ if $\mathcal{L}_1$ and $\mathcal{L}_2$ are linear networks with

$$\mathcal{L}_1 \cap \mathcal{L}_2 = \{u_0\} \quad \text{(a single point)}$$

$$d(u, v) = d(u, u_0) + d(v, u_0) \quad \text{whenever } u \in \mathcal{L}_1 \text{ and } v \in \mathcal{L}_2$$

(impossible if d is usual Euclidean distance).



Induction: $L = \mathcal{L}_1 \cup \ldots \cup \mathcal{L}_n$ is a 1-sum of $\mathcal{L}_1, \ldots, \mathcal{L}_n$ if L is a 1-sum of $\mathcal{L}_1 \cup \ldots \cup \mathcal{L}_{n-1}$ and $\mathcal{L}_n$.

Some main results

(Anderes, Møller & Rasmussen, 2022, Theorems 1 and 2)

Theorem

Let $r_0 : [0, \infty) \mapsto \mathbb{R}$ *be completely monotone and* $r_0 \neq 1$*: We have*

- $r_0(d_{\mathcal{R}}(u, v))$ *is strictly positive definite over* $(u, v) \in L \times L$*,*
- *if* L *is a 1-sum of trees and loops, then* $r_0(d_{\mathcal{G}}(u, v))$ *is strictly positive definite over* $(u, v) \in L \times L$*,*
- *if there are three distinct paths between two points on* L*, then there is some* $\phi > 0$ *so that* $\exp(-d_{\mathcal{G}}(u, v)/\phi)$ *is not positive definite over* $(u, v) \in L \times L$*.*

Exercise

Construct an example of a linear network which is forbidden for $d_{\mathcal{G}}$ (in the sense of the last statement in the theorem above).

Simulation of GPs on linear networks

Let $Y = \{Y(u) \mid u \in L\}$ be a mean zero GP on L.

A straightforward algorithm:

- select finite subsets $D_j \subset L_j$, $j = 1, \ldots, m$
- simulate Y restricted to $D = V \cup D_1 \ldots \cup D_m$ (e.g. using Choleski decomposition)
- for $u \notin D$, approximate $Y(u)$ by the average of those $Y(v)$ where $v \in D$ is closest to u ('closest' with respect to the given metric d or perhaps better $d_{\mathcal{G}}$).

Disadvantage: the dimension of Y can be large and hence Choleski decomposition (as well as other methods) can be slow.

A faster simulation algorithm for GPs on trees

First, we need some terminology: Let L be a tree, pick an arbitrary origin $u_0 \in V$, and set $G_0(u_0) = \{u_0\}$.

- Call $G_1(u_0) = \{v \in V : u_0 \sim v\}$ the first generation to u_0.
- Succeed for $j = 2, 3, \ldots$ to obtain the second, third, ... generations $G_2(u_0), G_3(u_0), \ldots$ to u_0.
- For the GP $Y = \{Y(w) \,|\, w \in L\}$ constructed in the following theorem, if $u \in G_{j-1}(u_0)$, $v \in G_j(u_0)$, and $u \sim v$, define

$$Y(u, v) = \{Y(w) : w \in (u, v]\}$$

where $(u, w]$ is the half-open line segment with endpoints u and v so that u is excluded and v is included.

A faster simulation algorithm (cont.)

Theorem

Suppose that L is a tree, and that $s > 0$ and $\sigma > 0$. Then $Y = \{Y(u) \mid u \in L\}$ constructed as follows is a mean-zero GP with exponential covariance function $c(u, v) = \sigma^2 \exp(-sd(u, v))$ where $d = d_{\mathcal{G}} = d_{\mathcal{R}}$.

- *For $w = u_0$, generate $Y(w) \sim N(0, \sigma^2)$.*
- *For $j = 1, 2, \ldots$, conditioned on all the $Y(w)$ so far generated, generate independent GPs $Y(u, v)$ for all $u \in G_{j-1}(u_0)$ and $v \in G_j(u_0)$ with $u \sim v$, where $Y(u, v)$ depends only on $Y(u)$ and for every $w, w_1, w_2 \in (u, v]$ we have*

$$\mathbb{E}[Y(w) \mid Y(u)] = \mathrm{e}^{-\|w-u\|s} Y(u)$$

$$\mathbb{C}\mathrm{ov}[Y(w_1, w_2) \mid Y(u)] = \sigma^2 \left(\mathrm{e}^{-\|w_1-w_2\|s} - \mathrm{e}^{-\|w_1-u\|s-\|w_2-u\|s}\right).$$

Simulation of CPs if c is not the exponential

Theorem

Suppose d is a metric on L so that $(u, v) \mapsto \exp(-sd(u, v))$ for $(u, v) \in L^2$ is a well-defined correlation function for all $s > 0$, and let Y be a mean-zero GP on L with covariance function

$$c(u, v) = \sigma^2 \mathbb{E} \exp(-d(u, v)S)$$

where $\sigma > 0$ and $S \geq 0$ is a random variable. For an integer $n > 0$ and $i = 1, \ldots, n$, generate a copy S_i of S and a mean-zero GPs Y_i on L with covariance function $\sigma^2 \exp(-S_i d(u, v))$ so that $(S_1, Y_1), \ldots, (S_n, Y_n)$ are independent. Calculate $\bar{Y}_n = \sum_{i=1}^n Y_i/n$. Then $\sqrt{n}\bar{Y}_n$ is a mean-zero stochastic process on L with covariance function c. As $n \to \infty$, $\sqrt{n}\bar{Y}_n$ approximates Y in the sense that any finite dimensional distribution of $\sqrt{n}\bar{Y}_n$ converges in distribution towards the corresponding finite dimensional distribution of Y.

Remarks

The theorem allows simulation of any GP with a covariance function of that form, provided a simulation algorithm for S is available and the metric d satisfies the condition in the theorem.

For the case of a tree and $d = d_{\mathcal{G}} = d_{\mathcal{R}}$, this gives a fast simulation algorithm when combined with the algorithm in the previous theorem.

For other cases it may be faster simply to use the straight forward algorithm, provided of course that d satisfies the condition in the theorem (so it may work for $d = d_{\mathcal{R}}$ but not for $d = d_{\mathcal{G}}$).

Example

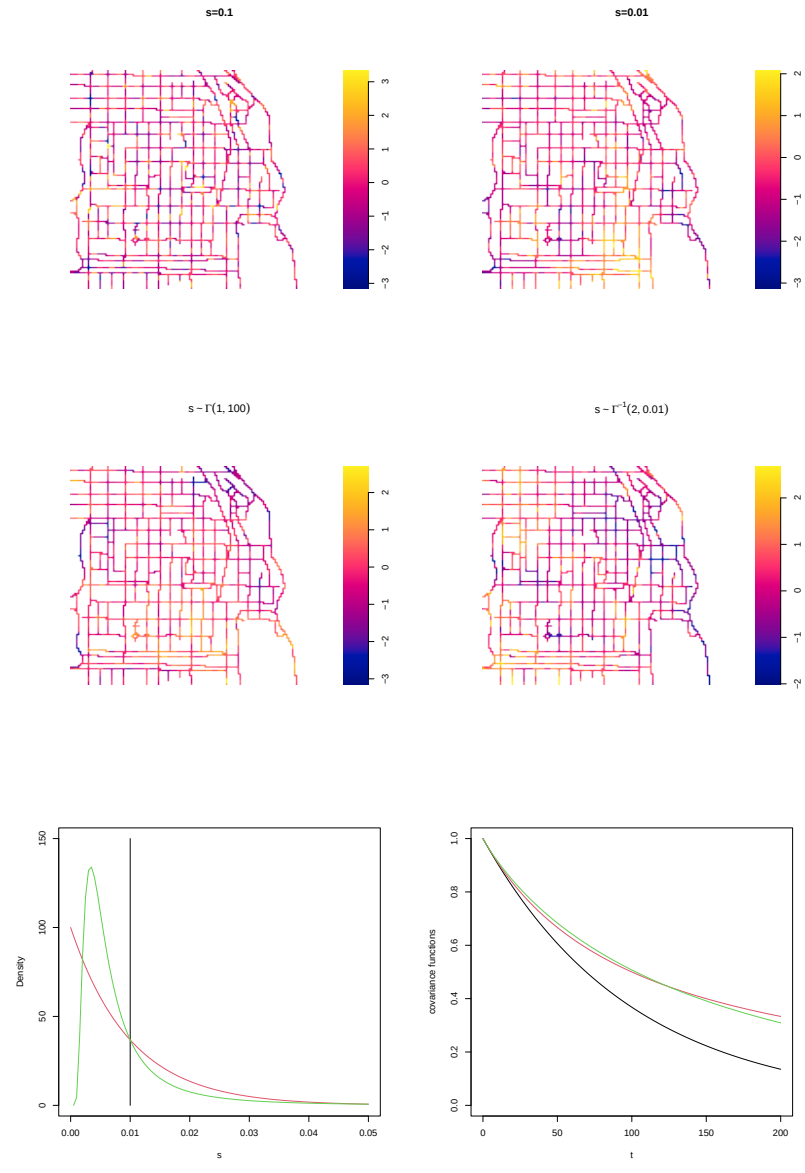


Figure 1: Simulation of zero mean GPs on the street network.

Details

Well-defined covariance functions $c(u, v) = r_0(d_{\mathcal{R}}(u, v))$ since the resistance metric has been used!

Top row: When $r_0(t) = \exp(-st)$ with parameter $s = 0.1$ or $s = 0.01$.

Middle row: When r_0 is a mixture of exponential correlation functions with s following a gamma distribution or inverse gamma distribution, where in both cases the mean of s is 0.01.

Bottom row: For plots 2–4, the corresponding pdf for S and correlation functions, where the curves in black, red, and green correspond to plots 2–4, respectively. NB: The side length of a square surrounding the network $\approx$ 1000 feet, but $\max d_{\mathcal{R}} \approx 675$ feet.

Point processes on L

In brief, a point process on L is a random finite subset $X \subset L$.

X is Poisson process with intensity function $\rho : L \mapsto [0, \infty)$ if

- $\#X \sim \text{Poisson}(\int \rho(u)\, \mathrm{d}_L(u))$
- the locations of the points in X do not depend on $\#X$ and they are i.i.d., each point having a density $\propto \rho$ (w.r.t. d_L).

Reduced Palm distribution

For $u \in L$, let $X_u \sim$ the reduced Palm distribution of X at u: Intuitively, X_u follows the distribution of $X \setminus \{u\}$ conditioned on that $u \in X$. (Formally:

$$\mathbb{E} \sum_{u \in X} h(X \setminus \{u\}, u) = \int \rho(u) \mathbb{E} h(X_u, u) \, \mathrm{d}_L(u)$$

for any non-negative (measurable) function h.)

If X has density f w.r.t. the unit rate Poisson process and $\rho(u) > 0$, then X_u has density

$$f_u(\{u_1, \ldots, u_n\}) = f(\{u, u_1, \ldots, u_n\})/\rho(u).$$

If X is Poisson process with intensity function ρ, then X and X_u are identically distributed (whenever $\rho(u) > 0$).

Moment and invariance properties

X has n-th order intensity function $\rho(u_1, \ldots, u_n)$ if

$$\mathbb{E}\left[N(A_1) \cdots N(A_n)\right] = \int_{A_1} \cdots \int_{A_n} \rho(u_1, \ldots, u_n)\, \mathrm{d}_L(u_1) \cdots \mathrm{d}_L(u_n)$$

is finite for any pairwise disjoint bounded $A_1, \ldots, A_n \subset L$, so $\rho(u)$ is the usual intensity function and

$$g(u, v) = \frac{\rho(u, v)}{\rho(u)\rho(v)}$$

is the pair correlation function (pcf).

We talk about stationary point processes on Euclidean spaces and isotropic point processes on spheres. In contrast, on a linear network there is no natural (transitive) group action.

One pays attention to isotropic pair correlation functions

$$g(u, v) = g_0(d(u, v)).$$

Properties of the pcf

For a Poisson process, $\rho(u_1, \ldots, u_n) = \rho(u_1) \cdots \rho(u_n)$, so $g = g_0 = 1$.

One often interprets

- $g(u, v) \leq 1$ as repulsion/inhibition/regularity
- $g(u, v) \geq 1$ as attraction/clustering/aggregation,

though care should be taken if u and v are 'distant apart'.

Non-parametric estimation

Non-parametric estimation of ρ and g_0: See Baddeley, Rubak & Turner (2015), Rakshit, Nair & Baddeley (2017) and Rakshit, Davies, Moradi, McSwiggan, Nair, Mateu & Baddeley (2019).

Re. the estimation of g_0, kernel methods are used but they may be sensible to the choice of bandwidth. Popular alternatives are given by estimators of the K-function below, however, since K is an accumulated versions of g_0, it may be harder to interpret.

For the specific parametric models in this presentation, we have simple expressions for g_0 but not for K. (Therefore, if time does not allow, the following 6 pages about K may be skipped!)

K-function (following Rakshit, Nair & Baddeley, 2017)

Suppose X is δ-correlated, i.e., $g(u, v) = g_0(\delta(u, v))$ is isotropic (we switch from d to δ since we consider derivatives below).

Let $R = \inf_{u \in L} \sup_{v \in L} \delta(u, v)$, then

$$K(t) = \int_0^t g_0(r)\,\mathrm{d}r, \quad 0 \leq t \leq R.$$

So K depends only on g_0, but g depends on both g_0 and δ.

If X is a Poisson process, then $K(t) = t$.

K-function (cont.)

In terms of Palm probabilities, for d_L-almost all $u \in L$ with $\rho(u) > 0$,

$$K(t) = \frac{1}{|L|}\mathbb{E}\sum_{v \in X_u} \frac{1(\delta(u,v) \leq t) w_\delta(u, \delta(u,v))}{\rho(v)},$$

where $|L| = \sum_{i=1}^{m} l_i$ and w_δ is a weight:

- w_δ is defined on the next slide,
- it accounts for the geometry of the linear network (when shifting from arc length measure on L to Lebesgue measure on the positive half-line, cf. Rakshit et al., 2017, Propositions 1 and 2),
- it makes it hard to interpret K.

Definition of the weight

Suppose δ is regular: for every $u \in L$, $\delta(u, v)$ is a continuous function of $v \in L$ and there is a finite set $N \subset L$ such that for $i = 1, \ldots, m$ and all $v \in L_i \setminus N$, the Jacobian

$$J_\delta(u, v) = |(\mathrm{d}/\mathrm{d}t)\delta(u, v)|$$

exists and is non-zero where $t = \|v - a_i\|$.

Def.: For $u \in L$ and $0 \leq t \leq R$,

$$1/w_\delta(u, t) = \sum_{v \in L:\, \delta(u,v)=t} 1/J_\delta(u, v).$$

Non-parametric estimation of K

Is carefully studied in Rakshit, Nair & Baddeley (2017) (still assuming δ is regular): for $A \subseteq L$ of positive arc length measure,

$$K(t) = \frac{1}{\int_A \mathrm{d}_L(u)} \mathbb{E} \sum_{u \in X_A} \sum_{v \in X \setminus \{u\}} \frac{1(\delta(u,v) \leq t) w_\delta(u, \delta(u,v))}{\rho(u)\rho(v)}$$

- Non-parametric estimators are based on omitting the expectation symbol and elaborating on the right hand side in order to realize how correction factors can be included in order to adjust for edge effects...
- Need to calculate w_δ!

The weight when $\delta = d_{\mathcal{G}}$

Disc w.r.t. $d_{\mathcal{G}}$ at radius $t = 300$ feet:

For $\delta = d_{\mathcal{G}}$: since $J_{d_{\mathcal{G}}} = 1$,

$$w_{\mathcal{G}}(u, t) = 1/\#\{v \in L \,|\, \delta(u, v) = t\}.$$

The weight when $\delta = d_{\mathcal{R}}$

For all $u \in L_j$ and $v \in L_i$ with $u \neq v$, $w_{d_{\mathcal{R}}}(u, v)$ is quickly calculated from its definition and our previous expression of $d_{\mathcal{R}}(u, v)$:

If $s = \|u - b_j\|$ and $t = \|v - a_i\|$ then $\frac{\mathrm{d}}{\mathrm{d}t} d_{\mathcal{R}}(u, v)$ is given by

$$\begin{cases} 2A_i(t-s) + 1 & \text{if } i = j,\ t > s, \\ 2A_i(t-s) - 1 & \text{if } i = j,\ t < s, \\ 2A_i t + B_{ij}(s) & \text{if } i \neq j. \end{cases}$$

Cox and Gaussian processes

In the following we construct Cox point processes generated by underlying GPs s.t. g is isotropic:

Let X be a Cox process driven by

$$\Lambda(u) = \rho(u)\Lambda_0(u), \quad u \in L,$$

where

- the intensity $\rho(u)$ may depend on covariate information
- the residual process Λ_0 accounts for unobserved covariates (Møller & Waagepetersen, 2007) and is specified by a transformation of one or more GPs s.t. $\mathbb{E}\Lambda_0(u) = 1$, $u \in L$.

Cox and Gaussian processes (cont.)

In general the likelihood is intractable but moment properties will be known:

$$\rho(u_1, \ldots, u_n) = \rho(u_1) \cdots \rho(u_n) \mathbb{E}\left[\Lambda_0(u_1) \cdots \Lambda_0(u_n)\right].$$

The models are typically attractive, i.e., $g \geq 1$.

Usually, we have even more:
For any pairwise disjoint bounded (Borel) sets $A_1, A_2, \ldots, A_n \subset L$,

$$\mathbb{E}\left[N(A_1) \cdots N(A_n)\right] \geq \mathbb{E}\left[N(A_1) \cdots N(A_i)\right] \mathbb{E}\left[N(A_{i+1}) \cdots N(A_n)\right]$$

(the counts are positively correlated at all orders).

Exercise

Verify the formulas on the previous slide and perhaps some of those for LGCPs, ICPs, and PPPs as given on the next slides.

Log Gaussian Cox process (LGCP)

Let $Y = \{Y(u) \,|\, u \in L\}$ be a GP with mean function μ and covariance function c.

X is a LGCP (Møller, Syversveen & Waagepetersen, 1998) if

$$\Lambda_0(u) = \exp(Y(u))$$

and $\mu(u) = -c(u,u)/2$ for all $u \in L$ (thus $\mathbb{E}\Lambda_0(u) = 1$).

- $\mathcal{D}(X)$ is determined by ρ and c or equivalently $g = \exp(c)$.
- g is isotropic iff c is isotropic.
- X_u is a LGCP with int. fct. $\rho(v|u) = \rho(v)\exp(c(u,v))$ and pcf $g(v,w|u) = g(v,w) = \exp(c(v,w))$ for $v, w \in L$ (Coeurjolly, Møller & Waagepetersen, 2017).

Interrupted Cox process (ICP)

Let $Y_1, \ldots, Y_h$ be i.i.d. zero mean GPs and $\Pi(u) = \exp(-\sum_{i=1}^{h} Y_i(u)^2)$.

X is an ICP (Stoyan, 1979; Lavancier & Møller, 2016) if

$$\Lambda_0(u) = \Pi(u)(1 + 2c(u,u))^{h/2}$$

for all $u \in L$, i.e., $\mathbb{E}\Lambda_0(u) = 1$ and $X|\Pi$ is an independent Π-thinning of a Poisson process on L with intensity function $u \mapsto \rho(u)\,(1 + c(u,u))^{h/2}$.

- $\mathcal{D}(X)$ is determined by (ρ, c, h).
- g is isotropic iff c is isotropic, in which case

$$g_0(t) = \left(\frac{(1+\sigma^2)^2}{(1+\sigma^2)^2 - \sigma^4 r_0(t)^2} \right)^{h/2}.$$

- As h increases, the degree of clustering decreases.

Permanental point process (PPP)

Let $Y_1, \ldots, Y_h$ be i.i.d. zero mean unit variance GPs.

X is a PPP (Macchi, 1975; McCullagh & Møller, 2006) if

$$\Lambda_0(u) = \frac{1}{h} \sum_{i=1}^{h} Y_i(u)^2$$

(more precisely X is a PPP with parameters $\alpha = h/2$ and $C(u, v) = \sqrt{\rho(u)\rho(v)}c(u, v)/\alpha$, and a Boson process if $\alpha = 1$).

- $\mathcal{D}(X)$ is determined by (ρ, c, h), and the degree of clustering is a decreasing function of α.
- $$g(u, v) = 1 + c(u, v)^2/\alpha \leq 1 + 1/\alpha \leq 3,$$

  showing the limitation of modelling clustering by a PPP.

Permanental point process (cont.)

- For $n = 1, 2, \ldots$ and $u_1, \ldots, u_n \in S$,

$$\rho(u_1, \ldots, u_n) = \rho(u_1) \cdots \rho(u_n) \mathrm{per}_\alpha[c](u_1, \ldots, u_n)/\alpha^n,$$

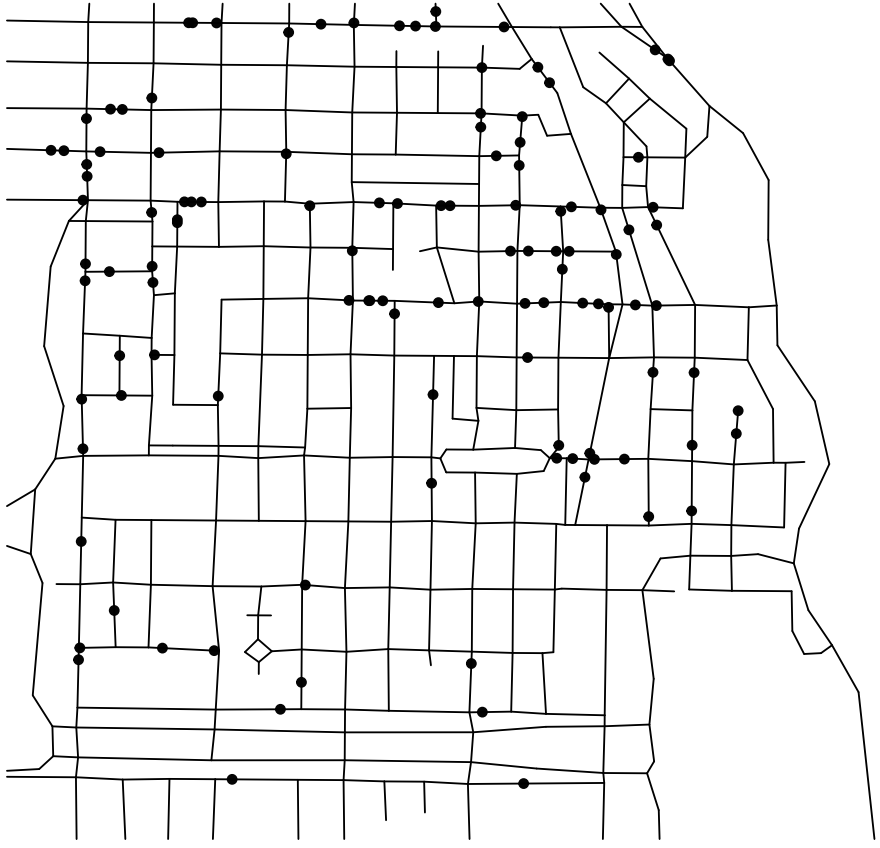
where we define the α-weighted permanent by

$$\mathrm{per}_\alpha[c](u_1, \ldots, u_n) = \sum_\pi \alpha^{\#\pi} c(u_1, u_{\pi_1}) \cdots c(u_n, u_{\pi_n})$$

where the sum is over all permutations $\pi = (\pi_1, \ldots, \pi_n)$ of $(1, \ldots, n)$ and $\#\pi$ is the number of cycles.

Data example 1 (Møller & Rasmussen, 2022)

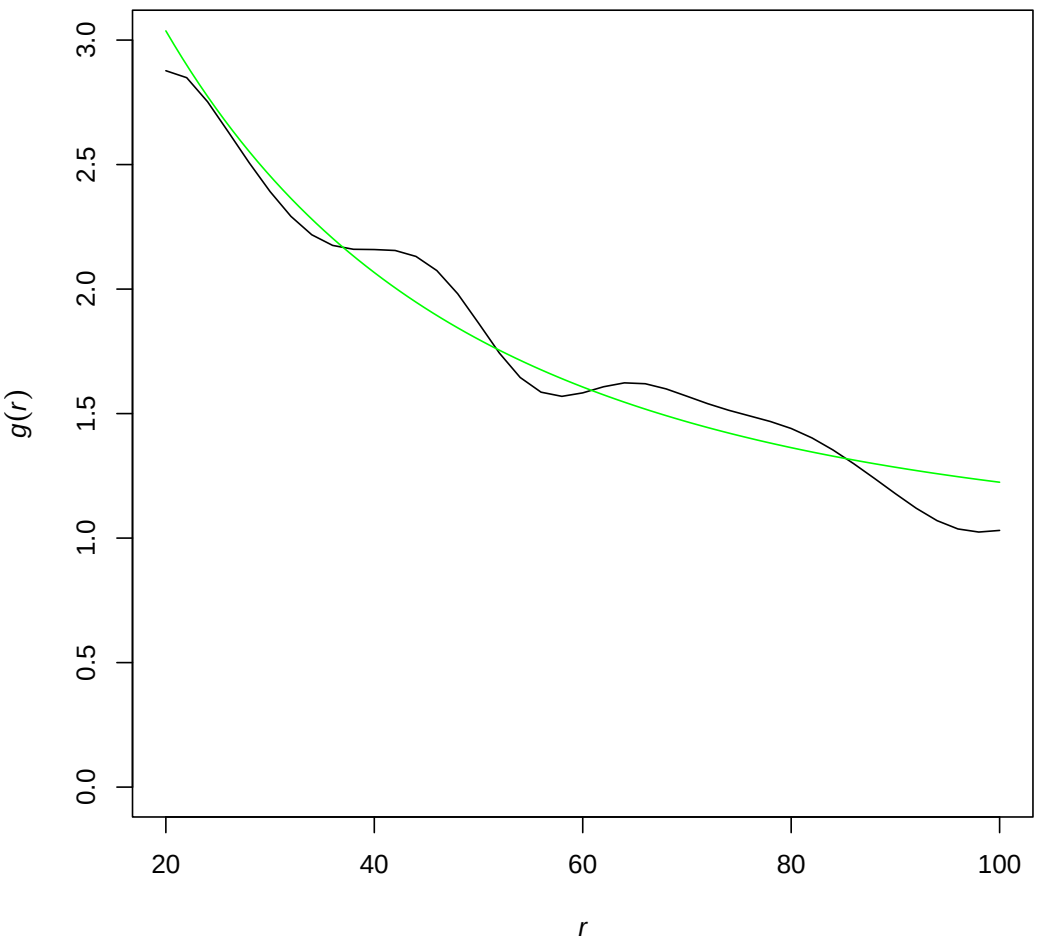
Crimes in a part of Chicago



Fitting a LGCP model with exp. cov. fct.

We need to use $d_{\mathcal{R}}$ in the exp. cov. fct.

Observed and fitted pcf (minimum contrast):



Model checking: Empirical F, G, J

These are three purely empirical summary functions:

- We skip the formal definitions (see Christensen & Møller, 2020), but obtained by modifying the empirical F-, G-, and J-functions for inhomogeneous point patterns on a Euclidean space (introduced by Van Lieshout, 2011) to linear networks.
- Do not correct for the network geometry, so their shapes alone can in general not be used to conclude anything about e.g. the presence of regularity or clustering.
- Still useful tools for providing a global rank envelope, i.e., a confidence region for a given test function obtained from simulations under a fitted model (Myllymäki, Mrkvička, Grabarnik, Seijo & Hahn, 2017).

Checking fitted LGCP

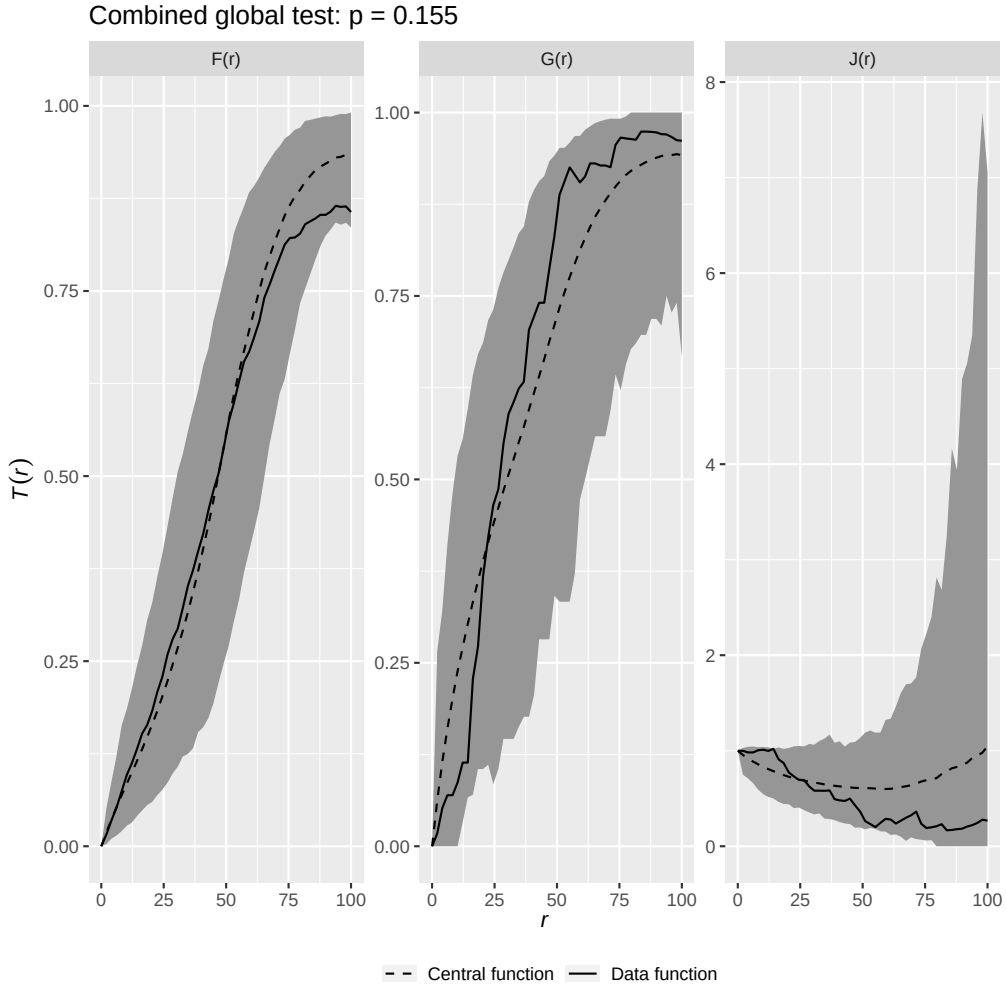
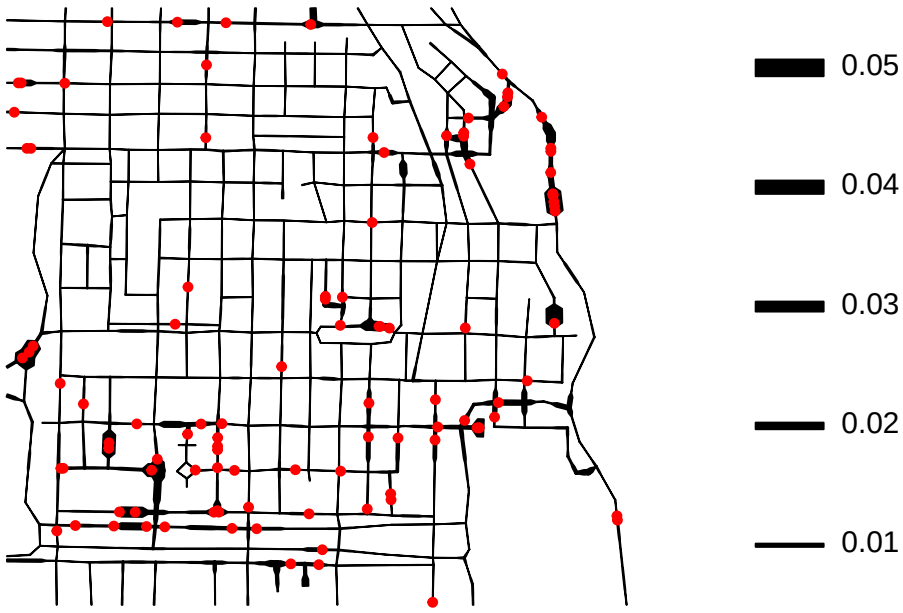


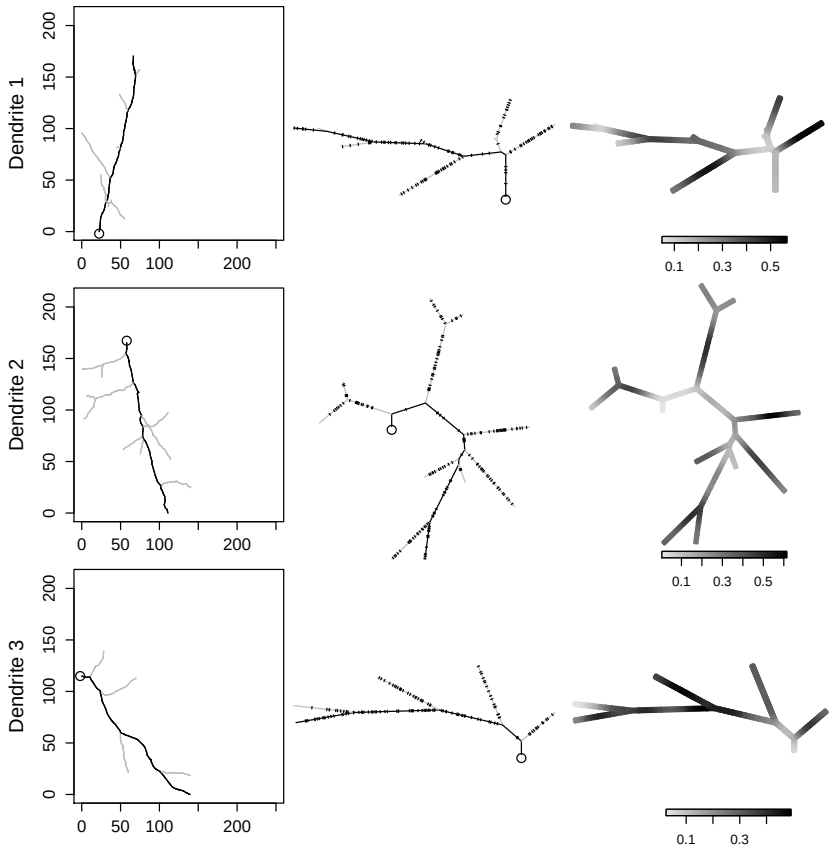
Figure 2: Concatenation of $\hat{F}$, $\hat{G}$, and $\hat{J}$ (black solid line) along with 95% global rank envelopes (grey region) and p-value for the associated global rank envelope test.

Simulation under fitted LGCP



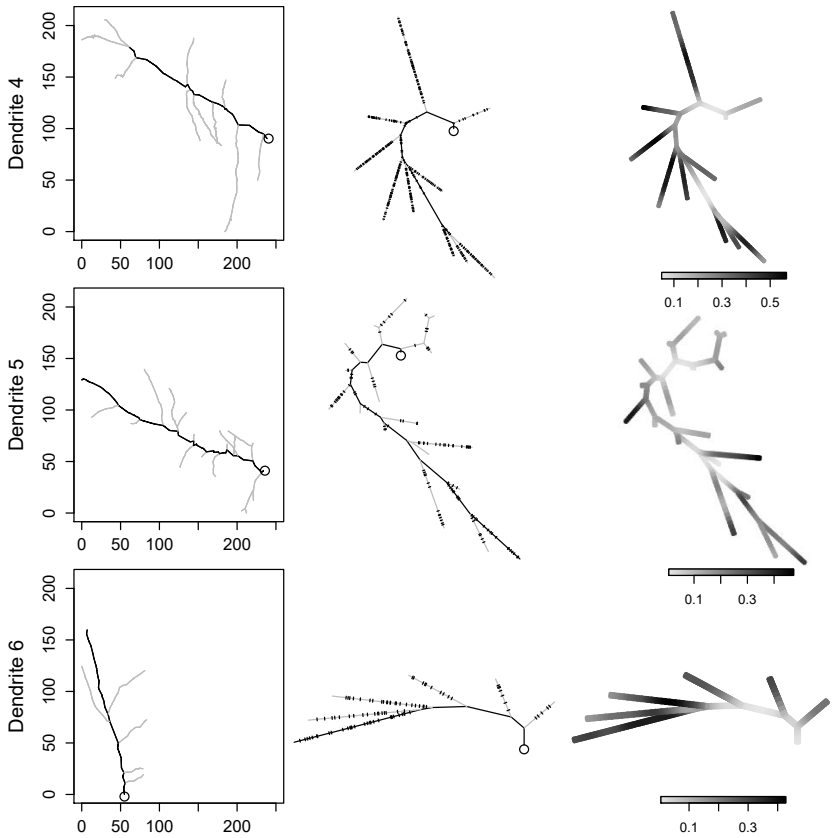
Data example 2 (Christensen & Møller, 2020)

Spine locations at first 3 dendrite trees: Planar projection (left; main branch is black), simplified networks embedded in $\mathbb{R}^2$ (middle; distances are preserved), and non-parametric kernel intensity estimates (right).



Data example (Christensen & Møller, 2020)

Spine locations at next 3 dendrite trees: Planar projection (left; main branch is black), simplified networks embedded in $\mathbb{R}^2$ (middle; distances are preserved), and non-parametric kernel intensity estimates (right).



Checking fitted inhomogeneous Poisson processes

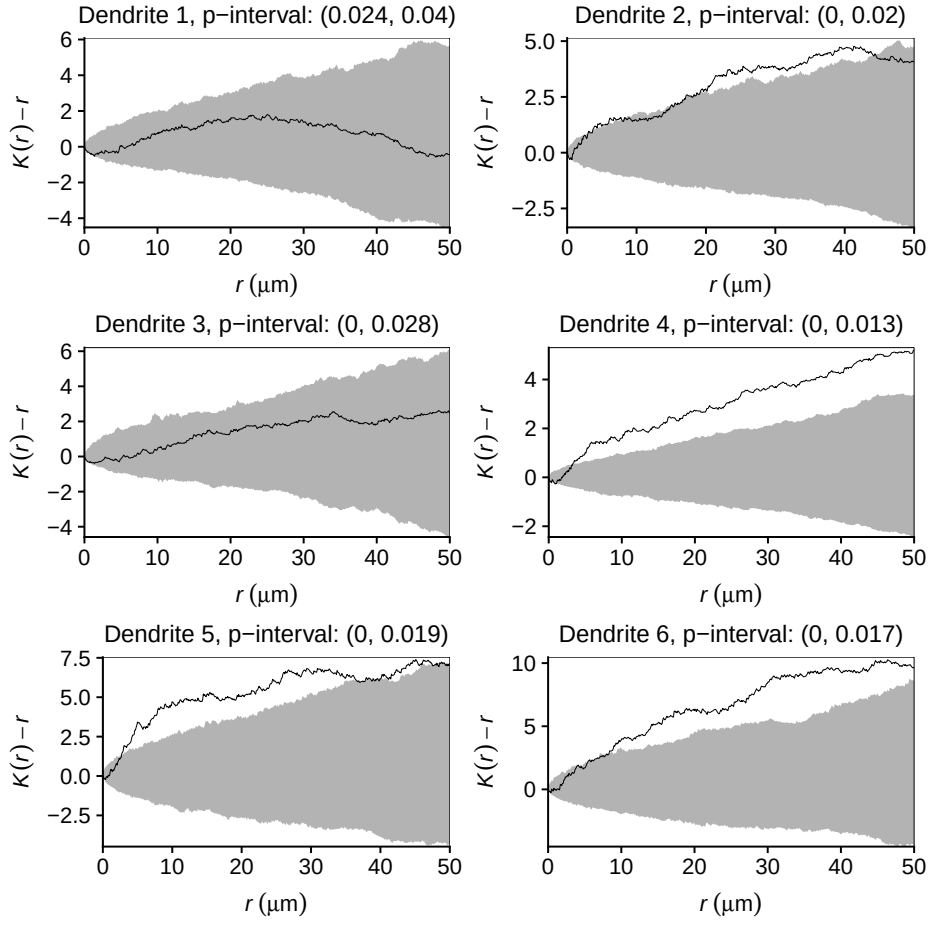


Figure 3: $\hat{K}$ minus the K-function for a Poisson process (solid line) along with 95% global rank envelopes (grey region) based on 2499 simulations from the fitted inhomogeneous Poisson process model, and p-intervals for the associated global rank envelope tests.

ICP model for each dendrite tree

Let

$$\rho(u) = \rho_1 1(u \in \text{main branch}) + \rho_2 1(u \in \text{side branch})$$

and

$-\ln \Pi =$ squared zero mean GP, $c(u, v) = \sigma^2 \exp(-\beta d_{\mathcal{G}}(u, v))$.

Estimate

- (ρ_1, ρ_2) by maximizing the first-order composite likelihood (Poisson likelihood)
- (σ, β) using a minimum contrast procedure based on g (performed better than using K or maximizing a second-order composite likelihood).

Checking fitted ICP models

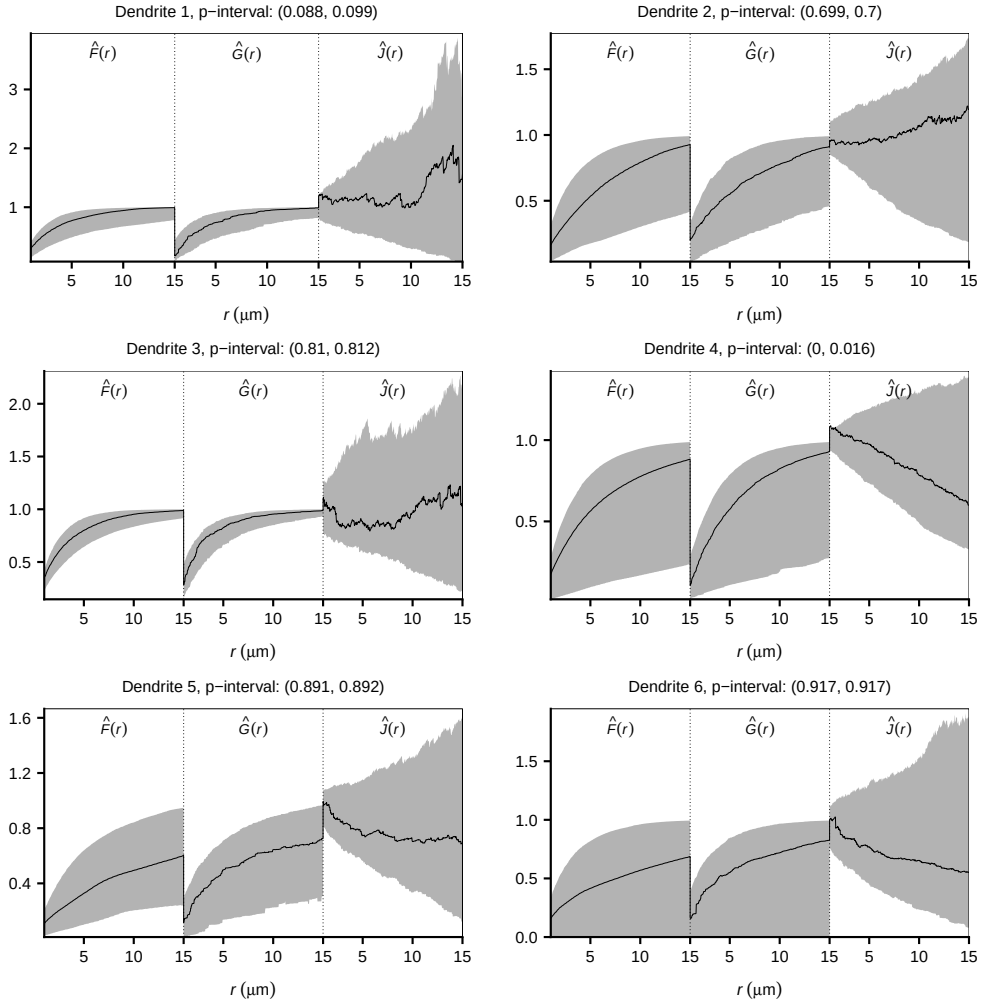


Figure 4: Concatenation of $\hat{F}$, $\hat{G}$, and $\hat{J}$ (black solid line) along with 95% global rank envelopes (grey region) based on 2499 simulations from the fitted ICP model; and p-intervals for the associated global rank envelope tests.

References

Anderes, E., Møller, J. and Rasmussen, J.G. (2020). Isotropic covariance functions on graphs and their edges. *Annals of Statistics*, **48**, 2478-2503.

Baddeley, A., Nair, G., Rakshit, S. and McSwiggan, G. (2017). "Stationary" point processes are uncommon on linear networks. *Stat*, **6**, 68-78.

Baddeley, A., Rubak, E. and Turner, R. (2015). Spatial Point Patterns: Methodology and Applications with R. Boca Raton: Chapman and Hall/CRC Press.

Christensen, H.S. and Møller, J. (2020). Modelling spine locations on dendrite trees using inhomogeneous Cox point processes. *Spatial Statistics*, **39**, Article 100478.

References (cont.)

Coeurjolly, J.-F., Møller, J. and Waagepetersen, R. (2017). Palm distributions for log Gaussian Cox processes. *Scandinavian Journal of Statistics*, **44**, 192–203.

Lavancier, F. and Møller, J. (2016). Modelling aggregation on the large scale and regularity on the small scale in spatial point pattern datasets. *Scandinavian Journal of Statistics*, **43**, 587-609.

Macchi, O. (1975). The coincidence approach to stochastic point processes. *Advances in Applied Probability*, **7**, 83-122.

McCullagh, P. and Møller, J. (2006). The permanental process. *Advances in Applied Probability*, **38**, 873-888.

References (cont.)

Møller, J. and J.G. Rasmussen (2022). Cox processes driven by transformed Gaussian processes on linear networks. *Scandinavian Journal of Statistics* (invited contribution; in preparation).

Møller, J., Syversveen, A.R. and Waagepetersen, R.P. (1998). Log Gaussian Cox processes. *Scandinavian Journal of Statistics*, **25**, 451-482.

Møller, J. and Waagepetersen, R. (2007). Modern statistics for spatial point processes (with discussion). *Scandinavian Journal of Statistics*, **34**, 643-711.

Myllymäki, M., Mrkvička, T., Grabarnik, P., Seijo, H. and Hahn, U. (2017). Global envelope tests for spatial processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **79**, 381-404.

References (cont.)

Okabe, A. and Sugihara, K. (2012). Spatial Analysis along Networks. New York: John Wiley and Sons.

Rakshit, S., Davies, T., Moradi, M., McSwiggan, G., Nair, G., Mateu, J. and Baddeley, A. (2019). Fast kernel smoothing of point patterns on a large network using two-dimensional convolution. *International Statistical Review*, **87**, 531-556.

Rakshit, S., Nair, G. and Baddeley, A. (2017). Second-order analysis of point patterns on a network using any distance metric. *Spatial Statistics*, **22**, 129-154.

Stoyan, D. (1979). Interrupted point processes. *Biometrical Journal*, **21**, 607-610.

Van Lieshout, M.N.M. (2011). A J-function for inhomogeneous point processes. *Statistica Neerlandica*, **65**, 183-201.