J-function for inhomogeneous spatio-temporal point processes

Citation for the original published paper (version of record):

Cronie, O., Van Lieshout, M. (2015). J-function for inhomogeneous spatio-temporal point processes. Scandinavian Journal of Statistics, 2: 562-579. http://dx.doi.org/10.1111/sjos.12123

Scandinavian Journal of Statistics Theory and Applications

Scandinavian Journal of Statistics, Vol. 42: 562–579, 2015
doi: 10.1111/sjos.12123

A *J*-function for Inhomogeneous Spatio-temporal Point Processes

O. CRONIE and M. N. M. VAN LIESHOUT
CWI, Amsterdam, The Netherlands

ABSTRACT. We propose a new summary statistic for inhomogeneous intensity-reweighted moment stationarity spatio-temporal point processes. The statistic is defined in terms of the n-point correlation functions of the point process, and it generalizes the J-function when stationarity is assumed. We show that our statistic can be represented in terms of the generating functional and that it is related to the spatio-temporal K-function. We further discuss its explicit form under some specific model assumptions and derive ratio-unbiased estimators. We finally illustrate the use of our statistic in practice.

Key words: inhomogeneous spatio-temporal point process, intensity-reweighted moment stationarity, J-function, K-function, location-dependent thinning of hard core model, log-Gaussian Cox process, n-point correlation function, Papangelou conditional intensity, Poisson process, (reduced Palm measure) generating functional

1. Introduction

A spatio-temporal point pattern can be described as a collection of pairs $\{(x_i, t_i)\}_{i=1}^m, m \geq 0$, where $x_i \in W_S \subseteq \mathbb{R}^d, d \geq 1$, and $t_i \in W_T \subseteq \mathbb{R}$ describe, respectively, the spatial location and the occurrence time associated with the i-th event. Examples of such point patterns include recordings of earthquakes, disease outbreaks and fires (see, e.g., Gabriel & Diggle (2009), Møller & Díaz-Avalos (2010) or Ogata (1998)).

When modelling spatio-temporal point patterns, the usual and natural approach is to assume that $\{(x_i, t_i)\}_{i=1}^m$ constitute a realization of a spatio-temporal point process (STPP) Y restricted to $W_S \times W_T$. Then, in order to deduce what type of model could describe the observations $\{(x_i, t_i)\}_{i=1}^m$, one carries out an exploratory analysis of the data under some minimal set of conditions on the underlying point process Y. At this stage, one is often interested in detecting tendencies for points to cluster together or to inhibit one another. In order to do so, one usually employs spatial or temporal summary statistics, which are able to capture and reflect such features.

A simple and convenient working assumption for the underlying point process is stationarity. In the case of a purely spatial point pattern $\{x_i\}_{i=1}^m \subseteq W_S$ generated by a stationary spatial point process X, a variety of summary statistics have been developed; see, for example, Chiu *et al.* (2013), Gelfand *et al.* (2010), Illian *et al.* (2008) or Lieshout (2000). One such statistic is the so-called J-function (Lieshout & Baddeley, 1996) given by

$$J(r) = \frac{1 - G(r)}{1 - F(r)} \qquad (1)$$

for $r \geq 0$ such that $F(r) \neq 1$. Here, the empty space function $F(r)$ is the probability of having at least one point of X within distance r from the origin, whereas the nearest neighbour distance distribution function $G(r)$ is the conditional probability of some further point of X falling within distance r from a typical point of X. Hence, $J(r) < 1$ indicates clustering, $J(r) = 1$ indicates spatial randomness and $J(r) > 1$ indicates regularity at inter-point distance r.

In many applications, though, stationarity is not a reasonable assumption. This observation has led to the development of summary statistics being able to compensate for inhomogeneity. For purely spatial point processes, Baddeley *et al.* (2000) introduced the notion of second-order intensity-reweighted stationarityand defined a summary statistic $K_{\text{inhom}}(r)$. It can be interpreted as an analogue of the K-function, which is proportional to the expected number of further points within distance r of a typical point of X, as it reduces to $K(r)$ when X is stationary.

The concept of second-order intensity-reweighted stationarity was extended to the spatio-temporal case by Gabriel & Diggle (2009) who also defined an inhomogeneous spatio-temporal K-function $K_{\text{inhom}}(r,t)$, $r,t \geq 0$. These ideas were further developed and studied in Møller & Ghorbani (2012) with particular attention to the notion of space–time separability.

To take into account interactions of order higher than two, Lieshout (2011) introduced the concept of intensity-reweighted moment stationarity for purely spatial point processes and generalized (1) to such point processes.

In this paper, we develop a proposal of Lieshout (2011) to study the spatio-temporal generalization $J_{\text{inhom}}(r,t)$ of (1) under suitable intensity reweighting. In section 2, we give the required preliminaries, which include definitions of product densities, Palm measures, generating functionals, n-point correlation functions and intensity-reweighted moment stationarity for STPPs. Then, in section 3, we give the definition of $J_{\text{inhom}}(r,t)$ under the assumption of intensity-reweighted moment stationarity and discuss its relation to the inhomogeneous spatio-temporal K-function of Gabriel & Diggle (2009). In section 4, we write $J_{\text{inhom}}(r,t)$ as a ratio of $1 - F_{\text{inhom}}(r,t)$ and $1 - G_{\text{inhom}}(r,t)$ in analogy with (1). As a by-product, we obtain generalizations of the empty space function and the nearest neighbour distance distribution. The section also includes a representation in terms of the Papangelou conditional intensity. In section 5, we consider three classes of STPPs for which the intensity-reweighted moment stationarity assumption holds, namely Poisson processes, location-dependent thinning of stationary STPPs and log-Gaussian Cox processes. In section 6, we derive a non-parametric estimator $\widehat{J_{\text{inhom}}}(r,t)$ for which we show ratio-unbiasedness. The new statistic is applied to data on the 2001 foot and mouth disease epidemic in the UK in section 7. Finally, section 8 provides a summary and discussion, and the Supporting Information contains some proofs and additional material.

2. Definitions and preliminaries

In this section, we recall notions from point process theory. The reader is referred to Daley & Vere-Jones (2003, 2008) for further details.

2.1. Simple spatio-temporal point process

In order to set the stage, let $\|x\| = \left(\sum_{i=1}^{d} x_i^2\right)^{1/2}$ and $d_{\mathbb{R}^d}(x,y) = \|x-y\|, x, y \in \mathbb{R}^d$, denote, respectively, the Euclidean norm and metric. Because space and time must be treated differently, we endow $\mathbb{R}^d \times \mathbb{R}$ with the supremum norm $\|(x,t)\|_\infty = \max\{\|x\|, |t|\}$ and the supremum metric

$$d((x,t),(y,s)) = \|(x,t)-(y,s)\|_\infty = \max\{d_{\mathbb{R}^d}(x,y), d_{\mathbb{R}}(t,s)\},$$

where $(x,t),(y,s) \in \mathbb{R}^d \times \mathbb{R}$. Then $(\mathbb{R}^d \times \mathbb{R}, d(\cdot,\cdot))$ is a complete separable metric space that is topologically equivalent to the Euclidean space $(\mathbb{R}^d \times \mathbb{R}, d_{\mathbb{R}^{d+1}}(\cdot,\cdot))$. Note that in the supremum metric, a closed ball of radius $r \geq 0$ centred at the origin $0 \in \mathbb{R}^d \times \mathbb{R}$ is given by the cylinder set

$$B[0,r] = \{(x,t) \in \mathbb{R}^d \times \mathbb{R} : \max\{\|x\|, |t|\} \leq r\}.$$

Write $\mathcal{B}(\mathbb{R}^d \times \mathbb{R}) = \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R})$ for the d-induced Borel σ-algebra, and let ℓ denote the Lebesgue measure on $\mathbb{R}^d \times \mathbb{R}$. Furthermore, given some Borel set $A \subseteq \mathbb{R}^d \times \mathbb{R}$ and some measurable function f, we interchangeably let $\int_A f(y)\,\mathrm{d}y$ and $\int_A f(y)\ell(\mathrm{d}y)$ represent the integral of f over A with respect to ℓ.

In this paper, an STPP is a simple point process on the product space $\mathbb{R}^d \times \mathbb{R}$. More formally, let N be the collection of all locally finite counting measures φ on $\mathcal{B}(\mathbb{R}^d \times \mathbb{R})$, that is, $\varphi(A) < \infty$ for bounded $A \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R})$, and let $\mathcal{N}$ be the smallest σ-algebra on N to make the mappings $\varphi \mapsto \varphi(A)$ measurable for all $A \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R})$. Consider in addition the sub-collection $N^* = \{\varphi \in N : \varphi(\{(x,t)\}) \in \{0,1\}$ for any $(x,t) \in \mathbb{R}^d \times \mathbb{R}\}$ of simple elements of N.

Definition 1. A *simple STPP* Y on $\mathbb{R}^d \times \mathbb{R}$ is a measurable mapping from some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ into the measurable space $(N, \mathcal{N})$ such that Y almost surely (a.s.) takes values in N^*.

Throughout, we will denote the Y-induced probability measure on $\mathcal{N}$ by P. Both $Y(\{(x,t)\}) = 1$ and $(x,t) \in Y$ will have the same meaning, and both $Y(A)$ and $|Y \cap A|$ may be used as notation for the number of points of Y in some set A, where $|\cdot|$ denotes cardinality.

2.2. *Product densities and n-point correlation functions*

Our definition of the inhomogeneous J-function relies on the so-called n-point correlation functions, which are closely related to the better known product densities. Here, we recall their definition.

Suppose that the factorial moment measures of Y exist as locally finite measures and that they are absolutely continuous with respect to the n-fold product of ℓ with itself. Then the Radon–Nikodym derivatives $\rho^{(n)}, n \geq 1$, referred to as *product densities*, are permutation invariant and defined by the integral equations

$$\begin{aligned} &\mathbb{E}\left[\sum^{\neq}_{(x_1,t_1),\ldots,(x_n,t_n)\in Y} h((x_1,t_1),\ldots,(x_n,t_n))\right] \\ &= \int \cdots \int h((x_1,t_1),\ldots,(x_n,t_n))\rho^{(n)}((x_1,t_1),\ldots,(x_n,t_n))\,\mathrm{d}x_1\,\mathrm{d}t_1 \cdots \mathrm{d}x_n\,\mathrm{d}t_n \end{aligned} \quad (2)$$

for non-negative measurable functions $h : (\mathbb{R}^d \times \mathbb{R})^n \to \mathbb{R}$ under the proviso that the left-hand side is infinite if and only if the right-hand side is. The symbol $\sum^{\neq}$ indicates that the sum is over n-tuples of distinct points. Equation (2) is sometimes referred to as the *Campbell formula*. Heuristically, $\rho^{(n)}((x_1,t_1),\ldots,(x_n,t_n))\,\mathrm{d}x_1\,\mathrm{d}t_1 \cdots \mathrm{d}x_n\,\mathrm{d}t_n$ is the infinitesimal probability of observing points $x_1,\ldots,x_n \in \mathbb{R}^d$ of Y at the respective event times $t_1,\ldots,t_n \in \mathbb{R}$. For $n = 1$, we obtain the *intensity measure* Λ of Y as

$$\Lambda(A) = \int_{B\times C} \rho^{(1)}((x,t))\,\mathrm{d}x\,\mathrm{d}t$$

for any $A = B \times C \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R})$. We shall also use the common notation $\lambda(x,t) = \rho^{(1)}((x,t))$ and assume henceforth that $\bar{\lambda} = \inf_{(x,t)} \lambda(x,t) > 0$.

The *n-point correlation functions* (White, 1979) are defined in terms of the $\rho^{(n)}$ by setting $\xi_1 \equiv 1$ and for other n recursively by

$$\frac{\rho^{(n)}((x_1,t_1),\ldots,(x_n,t_n))}{\prod_{k=1}^{n}\lambda(x_k,t_k)} = \sum_{k=1}^{n}\sum_{D_1,\ldots,D_k}\prod_{j=1}^{k}\xi_{|D_j|}(\{(x_i,t_i): i\in D_j\}), \tag{3}$$

where $\sum_{D_1,\ldots,D_k}$ is a sum over all possible k-sized partitions $\{D_1,\ldots,D_k\}$, $D_j \neq \emptyset$, of the set $\{1,\ldots,n\}$ and $|D_j|$ denotes the cardinality of D_j.

For a Poisson process on $\mathbb{R}^d \times \mathbb{R}$ with intensity function $\lambda(x,t)$, owing to, for example, theorem 1.3 in Lieshout (2000), $\rho^{(n)}((x_1,t_1),\ldots,(x_n,t_n)) = \prod_{k=1}^{n}\lambda(x_k,t_k)$, whereby $\xi_n \equiv 0$ for all $n \geq 2$. Hence, the sum on the right-hand side in expression (3) is a finite series expansion of the dependence correction factor by which we multiply the product density $\prod_{k=1}^{n}\lambda(x_k,t_k)$ of the Poisson process to obtain the product density $\rho^{(n)}((x_1,t_1),\ldots,(x_n,t_n))$ of Y.

A further interpretation is obtained by realizing that the right-hand side of (3) is a series expansion of a higher-order version of the *pair correlation function*

$$g((x_1,t_1),(x_2,t_2)) = \frac{\rho^{(2)}((x_1,t_1),(x_2,t_2))}{\lambda(x_1,t_1)\lambda(x_2,t_2)} = 1 + \xi_2((x_1,t_1),(x_2,t_2)).$$

The main definition of this section gives the class of STPPs to which we shall restrict ourselves in the sequel of this paper.

Definition 2. Let Y be an STPP for which product densities of all orders exist. If $\bar{\lambda} = \inf_{(x,t)}\lambda(x,t) > 0$ and, for all $n \geq 1$, ξ_n is translation invariant in the sense that

$$\xi_n((x_1,t_1)+(a,b),\ldots,(x_n,t_n)+(a,b)) = \xi_n((x_1,t_1),\ldots,(x_n,t_n))$$

for almost all $(x_1,t_1),\ldots,(x_n,t_n) \in \mathbb{R}^d \times \mathbb{R}$ and all $(a,b) \in \mathbb{R}^d \times \mathbb{R}$, we say that Y is intensity-reweighted moment stationary (IRMS).

By Equation (3), translation invariance of all ξ_n is equivalent to translation invariance of the intensity-reweighted product densities. The property is weaker than stationarity, which requires the distribution of Y to be invariant under translation, but stronger than the second order intensity-reweighted stationarity of Baddeley *et al.* (2000). The latter property, in addition to $\bar{\lambda} > 0$, requires the random measure

$$\Xi = \sum_{(x,t)\in Y}\frac{\delta_{(x,t)}}{\lambda(x,t)}$$

to be second order stationary (Daley & Vere-Jones, 2008, p. 236).

2.3. *Palm measures and conditional intensities*

In order to define a nearest neighbour distance distribution function, we need the concept of *reduced Palm measures*. In integral terms, recalling the assumption that the intensity measure is locally finite, they can be defined by the *reduced Campbell–Mecke* formula

$$\begin{aligned}\mathbb{E}\left[\sum_{(x,t)\in Y} g(x,t,Y\setminus\{(x,t)\})\right] &= \int_{\mathbb{R}^d\times\mathbb{R}}\int_N g(x,t,\varphi)P^{!(x,t)}(\mathrm{d}\varphi)\lambda(x,t)\,\mathrm{d}x\,\mathrm{d}t \\ &= \int_{\mathbb{R}^d\times\mathbb{R}} \mathbb{E}^{!(x,t)}\left[g(x,t,Y)\right]\lambda(x,t)\,\mathrm{d}x\,\mathrm{d}t\end{aligned} \tag{4}$$

for any non-negative measurable function $g: \mathbb{R}^d \times \mathbb{R} \times N \to \mathbb{R}$, with the left-hand side being infinite if and only if the right-hand side is infinite. By standard measure theoretic arguments

(Halmos, 1974), it is possible to find a regular version such that $P^{!(x,t)}(R)$ is measurable as a function of (x,t) and a probability measure as a function of R. Thus, $P^{!(x,t)}(R)$ may be interpreted as the conditional probability of $Y \setminus \{(x,t)\}$ falling in $R \in \mathcal{N}$ given $Y(\{(x,t)\}) > 0$.

At times, we make the further assumption that Y admits a *Papangelou conditional intensity* $\lambda(\cdot,\cdot;\varphi)$. In effect, we may then replace expectations under the reduced Palm distribution by expectations under P. More precisely, (4) may be rewritten as

$$\mathbb{E}\left[\sum_{(x,t)\in Y} g(x,t,Y\setminus\{(x,t)\})\right] = \int_{\mathbb{R}^d\times\mathbb{R}} \mathbb{E}\left[g(x,t,Y)\lambda(x,t;Y)\right] \mathrm{d}x\mathrm{d}t \tag{5}$$

for any non-negative measurable function $g \geq 0$ on $\mathbb{R}^d\times\mathbb{R}\times N$. Equation (5) is referred to as the *Georgii–Nguyen–Zessin* formula. We interpret $\lambda(x,t;Y)\,\mathrm{d}x\,\mathrm{d}t$ as the conditional probability of finding a space–time point of Y in the infinitesimal region $d(x,t)$ given that the configuration elsewhere coincides with Y.

2.4. The generating functional

For the representation of J_{inhom} in the form (1), we will need the *generating functional* $G(\cdot)$ of Y, which is defined as

$$G(v) = \mathbb{E}\left[\prod_{(x,t)\in Y} v(x,t)\right] = \int_N \prod_{(x,t)\in\varphi} v(x,t)\, P(\mathrm{d}\varphi)$$

for all functions $v = 1-u$ such that $u : \mathbb{R}^d \times \mathbb{R} \to [0,1]$ is measurable with bounded support on $\mathbb{R}^d \times \mathbb{R}$. By convention, an empty product equals 1. The generating functional uniquely determines the distribution of Y (theorem 9.4.V. in Daley & Vere-Jones (2008)).

Because we assume that the product densities of all orders exist,

$$\begin{aligned} G(v) &= G(1-u) = \\ &= 1 + \sum_{n=1}^{\infty} \frac{(-1)^n}{n!} \int \cdots \int u(x_1,t_1)\cdots u(x_n,t_n)\rho^{(n)}((x_1,t_1),\ldots,(x_n,t_n)) \prod_{i=1}^{n} \mathrm{d}x_i\,\mathrm{d}t_i, \end{aligned} \tag{6}$$

provided that the right-hand side converges (Chiu *et al.*, 2013, p. 126). The generating functional of the reduced Palm distribution $P^{!y}$ is denoted by $G^{!y}(v)$.

3. Spatio-temporal J-functions

We now turn to the definition of the inhomogeneous J-function. Before giving the definition in our general context, we define a spatio-temporal J-function for stationary STPPs.

3.1. The stationary J-function

Assume for the moment that Y is stationary. Then we may set, in complete analogy to the definition in Lieshout & Baddeley (1996),

$$J(r,t) = \frac{1-G(r,t)}{1-F(r,t)} = \frac{\mathbb{P}^{!(0,0)}\left(Y \cap S_r^t = \emptyset\right)}{\mathbb{P}\left(Y \cap S_r^t = \emptyset\right)} \tag{7}$$

for $r, t \geq 0$ such that $F(r,t) \neq 1$, where

$$S_r^t = \left\{(x,s) \in \mathbb{R}^d \times \mathbb{R} : \|x\| \leq r, |s| \leq t\right\}$$

and $\mathbb{P}^{!(0,0)}$ is the $P^{!(0,0)}$-reversely induced probability measure on $\mathcal{F}$.

Note that the two equalities in (7) are defining ones, $G(r,t)$ is the spatio-temporal nearest neighbour distance distribution function, and $F(r,t)$ the spatio-temporal empty space function.

3.2. *The inhomogeneous J-function*

In this section, we extend the inhomogeneous J-function of Lieshout (2011) to the product space $\mathbb{R}^d \times \mathbb{R}$ equipped with the supremum metric $d(\cdot,\cdot)$.

Definition 3. Let Y be an IRMS STPP (cf. definition 2). For $r, t \geq 0$, let

$$J_n(r,t) = \int_{S_r^t} \cdots \int_{S_r^t} \xi_{n+1}((0,0),(x_1,t_1),\ldots,(x_n,t_n)) \prod_{i=1}^{n} \mathrm{d}x_i \mathrm{d}t_i$$

and set

$$J_{\mathrm{inhom}}(r,t) = 1 + \sum_{n=1}^{\infty} \frac{\left(-\bar{\lambda}\right)^n}{n!} J_n(r,t) \tag{8}$$

for all spatial ranges $r \geq 0$ and temporal ranges $t \geq 0$ for which the series is absolutely convergent.

Note that by Cauchy's root test, absolute convergence holds for those $r, t \geq 0$ for which $\limsup_{n\to\infty} \left(\bar{\lambda}^n |J_n(r,t)|/n!\right)^{1/n} < 1$.

Let us briefly mention a few special cases. For a Poisson process, as $\xi_{n+1} \equiv 0$ for $n \geq 1$, $J_{\mathrm{inhom}}(r,t) \equiv 1$. Moreover, if Y is stationary, (8) reduces to (7).

3.3. *Relationship to K-functions*

The spatio-temporal K-function may be seen as a second order approximation of J_{inhom}. Indeed, Gabriel & Diggle (2009), under the assumptions that the intensity function is bounded away from zero and the pair correlation function $g((x,t),(y,s)) = \bar{g}(u,v)$ depends only on $u = \|x-y\|$ and $v = |t-s|$, introduce a spatio-temporal inhomogeneous K-function by setting

$$K_{\mathrm{inhom}}(r,t) = \int_{S_r^t} \bar{g}(\|x_1\|,|t_1|)\, \mathrm{d}(x_1,t_1) = \omega_d \int_{-t}^{t} \int_0^r \bar{g}(u,v) u^{d-1}\, \mathrm{d}u \mathrm{d}v,$$

where $\omega_d/d = \pi^{d/2}/\Gamma(1+d/2) = \kappa_d$ is the volume of the unit ball in $\mathbb{R}^d$ (see, e.g., Chiu *et al.* (2013, p. 14)). Note that the second equality follows from a change to hyperspherical coordinates. If in addition Y is IRMS,

$$\begin{aligned} J_{\mathrm{inhom}}(r,t) - 1 &= -\bar{\lambda}\left(\omega_d \int_{-t}^{t}\int_0^r \bar{g}(u,v)u^{d-1}\,\mathrm{d}u\mathrm{d}v - \ell\left(S_r^t\right)\right) + \sum_{n=2}^{\infty} \frac{\left(-\bar{\lambda}\right)^n}{n!} J_n(r,t) \\ &\approx -\bar{\lambda}\left(K_{\mathrm{inhom}}(r,t) - \ell\left(S_r^t\right)\right), \end{aligned}$$

whereby $K_{\mathrm{inhom}}(r,t)$ may be viewed as a second order counterpart of $J_{\mathrm{inhom}}(r,t)$. In relation hereto, it should be noted that even if the product densities exist only up to some finite order m, we may still obtain an approximation of J_{inhom} by truncating its series representation at $n = m$.

Under the assumption of intensity-reweighted moment stationarity that is, $\bar{\lambda} > 0$ and Ξ second-order stationary, we may extend the definition in Baddeley *et al.* (2000) to the spatio-temporal setting by defining

$$K^*_{\mathrm{inhom}}(r,t) = \frac{1}{\ell(A)}\mathbb{E}\left[\sum^{\neq}_{(x_1,t_1),(x_2,t_2)\in Y} \frac{\mathbf{1}\{(x_1,t_1)\in A, \|x_1-x_2\|\le r, |t_1-t_2|\le t\}}{\lambda(x_1,t_1)\lambda(x_2,t_2)}\right]$$

for $r,t \ge 0$ and some set $A = B \times C \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R})$ with $\ell(A) > 0$. By lemma 1 below, the definition does not depend on the choice of A.

Lemma 1. *For any $A = B \times C \in \mathcal{B}(\mathbb{R}^d \times \mathbb{R})$ for which $\ell(A) > 0$, $K^*_{\mathrm{inhom}}(r,t) = \mathcal{K}_\Xi\left(S_r^t \setminus \{(0,0)\}\right)$, the reduced second factorial moment measure of Ξ evaluated at S_r^t (see, e.g., section 12.6 in Daley & Vere-Jones (2008)).*

Proof. By the Campbell formula, the intensity measure of Ξ is given by

$$\Lambda_\Xi(A) = \mathbb{E}\left[\sum_{(x,t)\in Y}\frac{1}{\lambda(x,t)}\mathbf{1}_A(x,t)\right] = \ell(A),$$

so it is locally finite and has density 1. Hence, by proposition 13.1.IV of Daley & Vere-Jones (2008), there exist reduced Palm measures $P_\Xi^{!y_1}(R)$, $y_1 \in \mathbb{R}^d \times \mathbb{R}$, $R \in \mathcal{N}$, such that

$$\begin{aligned} K^*_{\mathrm{inhom}}(r,t) &= \frac{1}{\ell(A)}\mathbb{E}\left[\int_{\mathbb{R}^d\times\mathbb{R}} \mathbf{1}_A(y_1)\,\Xi\left((y_1+S_r^t)\setminus\{y_1\}\right)\,\Xi(\mathrm{d}y_1)\right] \\ &= \frac{1}{\ell(A)}\int_{\mathbb{R}^d\times\mathbb{R}} \mathbb{E}^{!y}\left[1_A(y)\,\Xi\left(y+S_r^t\right)\right]\mathrm{d}y = \mathbb{E}^{!(0,0)}\left[\Xi\left(S_r^t\right)\right] = \mathcal{K}_\Xi\left(S_r^t\right) \end{aligned}$$

and this completes the proof. □

It is not hard to see that under the stronger assumptions of Gabriel & Diggle (2009), $K^*_{\mathrm{inhom}}(r,t) = K_{\mathrm{inhom}}(r,t)$.

4. Representation results

Being based on a series of integrals of n-point correlation functions, definition 3 highlights the fact that J_{inhom} involves interactions of all orders but it is not very convenient in practice. The goal of this section is to give representations that are easier to interpret.

4.1. Representation in terms of generating functionals

As for purely spatial point processes, we may express J_{inhom} in terms of the generating functionals G and $G^{!\cdot}$ by an appropriate choice of $v = 1 - u$ (Lieshout, 2011). Indeed, set

$$u^y_{r,t}(x,s) = \frac{\bar{\lambda}\mathbf{1}\{\|a-x\|\le r, |b-s|\le t\}}{\lambda(x,s)}, \quad y=(a,b)\in\mathbb{R}^d\times\mathbb{R},$$

and define the *inhomogeneous spatio-temporal nearest neighbour distance distribution function* as

$$G_{\text{inhom}}(r,t) = 1 - G^{!y}\left(1 - u^{y}_{r,t}\right)$$
$$= 1 - \mathbb{E}^{!(a,b)}\left[\prod_{(x,s)\in Y}\left(1 - \frac{\bar{\lambda}\mathbf{1}\{\|a-x\| \le r, |b-s| \le t\}}{\lambda(x,s)}\right)\right] \tag{9}$$

and the *inhomogeneous spatio-temporal empty space function* as

$$F_{\text{inhom}}(r,t) = 1 - G\left(1 - u^{y}_{r,t}\right)$$
$$= 1 - \mathbb{E}\left[\prod_{(x,s)\in Y}\left(1 - \frac{\bar{\lambda}\mathbf{1}\{\|a-x\| \le r, |b-s| \le t\}}{\lambda(x,s)}\right)\right]$$

for $r, t \ge 0$, under the convention that empty products take the value one. Then, the representation theorem below tells us that $G_{\text{inhom}}(r,t)$ and $F_{\text{inhom}}(r,t)$ do not depend on the choice of y and, furthermore, that $J_{\text{inhom}}(r,t)$ may be expressed through $G_{\text{inhom}}(r,t)$ and $F_{\text{inhom}}(r,t)$. The proof of theorem 1 can be found in the Supporting Information (Appendix S1).

Theorem 1. *Let Y be an IRMS STPP and assume that*

$$\limsup_{n\to\infty}\left(\frac{\bar{\lambda}^n}{n!}\int_{S^t_r}\cdots\int_{S^t_r}\frac{\rho^{(n)}((x_1,t_1),\ldots,(x_n,t_n))}{\lambda(x_1,t_1)\cdots\lambda(x_n,t_n)}\prod_{i=1}^{n}\mathrm{d}x_i\,\mathrm{d}t_i\right)^{1/n} < 1.$$

Then $G_{\text{inhom}}(r,t)$ and $F_{\text{inhom}}(r,t)$ are ℓ-almost everywhere constant with respect to $y = (a,b) \in \mathbb{R}^d \times \mathbb{R}$ and the J-function of definition 3 can be written as

$$J_{\text{inhom}}(r,t) = \frac{1 - G_{\text{inhom}}(r,t)}{1 - F_{\text{inhom}}(r,t)}$$

for all $r, t \ge 0$ such that $F_{\text{inhom}}(r,t) \ne 1$.

The intuition behind $G_{\text{inhom}}(r,t)$ and $F_{\text{inhom}}(r,t)$ is best seen when Y is stationary. In this case, $u^0_{r,t}(x,s) = \mathbf{1}\left\{(x,s) \in S^t_r\right\}$, and hence

$$F_{\text{inhom}}(r,t) = 1 - \mathbb{E}\left[\prod_{(x,s)\in Y}\mathbf{1}\left\{(x,s) \notin S^t_r\right\}\right] = 1 - \mathbb{P}\left(Y \cap S^t_r = \emptyset\right) = F(r,t),$$

the empty space function in expression (7). Similarly, $G_{\text{inhom}}(r,t)$ reduces to the distribution function of the nearest neighbour distance when Y is stationary, and J_{inhom} is indeed a generalization of (1).

4.2. *Representation in terms of conditional intensities*

Some families of point processes, notably Gibbsian ones (Lieshout, 2000), are defined in terms of their Papangelou conditional intensity $\lambda(\cdot,\cdot;\cdot)$. In theorem 2, we show that for such processes, J_{inhom} may be represented in terms of $\lambda(\cdot,\cdot;\cdot)$. The proof of the theorem is given in the Supporting Information (Appendix S2).

Theorem 2. *Let the assumptions of theorem 1 hold and assume, in addition, that Y admits a conditional intensity $\lambda(\cdot,\cdot;\cdot)$. Write $W_{(a,b)}(Y) = \prod_{(x,s)\in Y}\left(1 - u^{(a,b)}_{r,t}(x,s)\right)$. Then $\mathbb{E}\left[\lambda(a,b;Y)W_{(a,b)}(Y)/\lambda(a,b)\right] > 0$ implies $E[W_{(a,b)}(Y)] > 0$ and*

$$J_{\text{inhom}}(r,t) = \mathbb{E}\left[\frac{\lambda(a,b;Y)}{\lambda(a,b)} W_{(a,b)}(Y)\right] / \mathbb{E}[W_{(a,b)}(Y)]$$

for almost all $(a,b) \in \mathbb{R}^d \times \mathbb{R}$.

Because $\mathbb{E}[\lambda(a,b;Y)] = \lambda(a,b)$, we immediately see that

$$J_{\text{inhom}}(r,t) \geq 1 \iff \text{Cov}\left(\lambda(a,b;Y), W_{(a,b)}(Y)\right) \geq 0$$

and

$$J_{\text{inhom}}(r,t) \leq 1 \iff \text{Cov}\left(\lambda(a,b;Y), W_{(a,b)}(Y)\right) \leq 0.$$

In words, for clustered point processes, $\lambda(a,b;Y)$ tends to be large if (a,b) is near to points of Y, whereas $W_{(a,b)}(Y)$ tends to be large when there are few points of Y close to (a,b). Thus, in this case, the two random variables are negatively correlated, and the J-function is smaller than one. A dual reasoning applies for regular point processes, but Bedford & Van den Berg (1997) warn against drawing too strong conclusions.

4.3. Scaling

In expression (8), we consider distances on the spaces $\mathbb{R}^d$ and $\mathbb{R}$ separately. Instead, we could have used the supremum distance on $\mathbb{R}^d \times \mathbb{R}$ and the closed d-metric balls $B[0,r] = S_r^r$, $r \geq 0$, to define $J_n(r) = J_n(r,r)$ and

$$J_{\text{inhom}}(r) = 1 + \sum_{n=1}^{\infty} \frac{(-\bar{\lambda})^n}{n!} J_n(r). \tag{10}$$

When the pair correlation function only depends on the spatial and temporal distances, set $K^*_{\text{inhom}}(r) = K^*_{\text{inhom}}(r,r)$ and $K_{\text{inhom}}(r) = K_{\text{inhom}}(r,r)$, whence $K^*_{\text{inhom}}(r) = K_{\text{inhom}}(r)$ and $J_{\text{inhom}}(r) - 1 \approx -\bar{\lambda}(K_{\text{inhom}}(r) - \ell(B[0,r]))$.

In the remainder of this section, we argue that (8) may be obtained from (10) by scaling. Let $c = (c_S, c_T) \in (0,\infty)^2$ and apply the bijective transformation $(y,s) \mapsto (c_S y, c_T s)$ to each point of the IRMS STPP Y to obtain

$$cY = \sum_{(y,s)\in Y} \delta_{(c_S y, c_T s)}.$$

Through a change of variables and the Campbell formula, one obtains

$$\rho^{(n)}_{cY}((x_1,t_1),\ldots,(x_n,t_n)) = c_S^{-dn} c_T^{-n} \rho^{(n)}((x_1/c_S, t_1/c_T),\ldots,(x_n/c_S, t_n/c_T)),$$

so that $\lambda_{cY}(x,t) = c_S^{-d} c_T^{-1} \lambda(x/c_S, t/c_T)$ and $\bar{\lambda}_{cY} = \inf_{(x,t)} \lambda_{cY}(x,t) = c_S^{-d} c_T^{-1} \bar{\lambda}$. Hence,

$$\xi_n^{cY}((x_1,t_1),\ldots,(x_n,t_n)) = \xi_n((x_1/c_S, t_1/c_T),\ldots,(x_n/c_S, t_n/c_T)),$$

whence cY is IRMS if and only if Y is and, whenever well defined,

$$J^{cY}_{\text{inhom}}(r,t) = J_{\text{inhom}}\left(\frac{r}{c_S}, \frac{t}{c_T}\right).$$

In conclusion, by taking $c_S = 1$, and $c_T = r/t$, any $J_{\text{inhom}}(r,t)$ may be obtained from $J_{\text{inhom}}(r)$ through scaling.

5. Theoretical examples

Next, we will consider three families of models, each representing a different type of interaction.

5.1. *Poisson processes*

The inhomogeneous Poisson process may be considered the benchmark for lack of interaction between points. As we saw in section 3.2, for a Poisson process $J_{\text{inhom}}(r,t) \equiv 1$. Alternative proofs may be obtained from theorems 1–2, by noting that the Palm distributions equal P by Slivnyak's theorem (Schneider & Weil, 2008) or that the intensity function and the Papangelou conditional intensity coincide almost everywhere.

5.2. *Location-dependent thinning*

Given a stationary STPP Y with product densities $\rho^{(n)}, n \geq 1$, intensity $\lambda > 0$ and J-function $J(r,t)$, consider some measurable function $p : \mathbb{R}^d \times \mathbb{R} \to (0,1]$ with $\bar{p} = \inf_{(x,t)} p(x,t) > 0$. Location-dependent thinning of Y is the scenario in which a point $(x,t) \in Y$ is retained with probability $p(x,t)$. Denote the resulting thinned process by Y_{th}. The product densities of Y_{th} are

$$\rho_{\text{th}}^{(n)}((x_1,t_1),\ldots,(x_n,t_n)) = \rho^{(n)}((x_1,t_1),\ldots,(x_n,t_n)) \prod_{i=1}^{n} p(x_i,t_i)$$

by (Daley & Vere-Jones, 2008, section 11.3), whereby $\lambda_{\text{th}}(x,t) = \lambda p(x,t) > 0$ and the n-point correlation functions of Y_{th} and Y coincide. Hence, Y_{th} is IRMS with $\bar{\lambda} = \inf_{(x,t)} \lambda_{\text{th}}(x,t) = \lambda \bar{p}$ and

$$J_{\text{inhom}}^{\text{th}}(r,t) = 1 + \sum_{n=1}^{\infty} \frac{(-\lambda\bar{p})^n}{n!} J_n(r,t)$$

for all $r,t \geq 0$ for which the series converges. Here, $J_n(r,t)$ is the n-th coefficient in the series expansion (8) of the J-function of the original process Y.

A more informative expression for $J_{\text{inhom}}^{\text{th}}$ can be obtained by noting that, by equations (5.3)–(5.4) in Chiu *et al.* (2013), the generating functional of Y_{th} is given by $G_{\text{th}}(v) = G(1 - p + pv)$, where $G(\cdot)$ is the generating functional of Y.

Hence, as applying thinning to the reduced Palm distribution of Y is equivalent to Palm conditioning in the thinned process,

$$J_{\text{inhom}}^{\text{th}}(r,t) = \frac{G_{\text{th}}^{!(0,0)}\left(1 - \bar{p}\mathbf{1}\{\cdot \in S_r^t\}/p(\cdot)\right)}{G_{\text{th}}\left(1 - \bar{p}\mathbf{1}\{\cdot \in S_r^t\}/p(\cdot)\right)} = \frac{\mathbb{E}^{!(0,0)}\left[(1-\bar{p})^{Y(S_r^t)}\right]}{\mathbb{E}\left[(1-\bar{p})^{Y(S_r^t)}\right]}$$

when theorem 1 applies.

When a Papangelou conditional intensity exists for Y, by recalling that $\lambda = \lambda(x,t) = \mathbb{E}[\lambda(x,t;Y)]$ and applying the combination of (4) and (5) to the restriction of the function $g(a,b,Y) = (1-\bar{p})^{Y((a,b)+S_r^t)}$ to arbitrary bounded space–time domains, the previous expression becomes

$$J_{\mathrm{inhom}}^{\mathrm{th}}(r,t) = \frac{\mathbb{E}\left[\lambda(0,0;Y)\,(1-\bar{p})^{Y(S_r^t)}\right]}{\lambda\mathbb{E}\left[(1-\bar{p})^{Y(S_r^t)}\right]}.$$

Thinned hard core process. The spatio-temporal hard core process is a stationary STPP defined through its Papangelou conditional intensity

$$\lambda_Y(a,b;Y) = \beta\,\mathbf{1}\left\{Y \cap \left((a,b) + S_{R_S}^{R_T}\right) = \emptyset\right\} = \beta \prod_{(x,t)\in Y} \mathbf{1}\left\{(x,t)-(a,b) \notin S_{R_S}^{R_T}\right\}, \tag{11}$$

where $(a,b) \in \mathbb{R}^d \times \mathbb{R}$. Moreover, $\beta > 0$ is a model parameter, and $R_S > 0$ and $R_T > 0$ are, respectively, the spatial and temporal hard core distances. In words, as realizations a.s. do not contain points that violate the spatial and temporal hard core constraints, that is, $\mathbb{P}^{!(0,0)}\left(Y\left(S_{R_S}^{R_T}\right) > 0\right) = 0$, there is inhibition.

By thinning Y with some suitable measurable retention function $p : \mathbb{R}^d \times \mathbb{R} \to (0,1]$, $\bar{p} = \inf_{(x,t)} p(x,t) > 0$, we obtain an IRMS hard core STPP Y_{th}.

Lemma 2. *For a hard core process $Y, \beta/\lambda \geq 1$. If either $(r,t) \in [0,R_S]\times[0,R_T]$ or $(r,t) \in [R_S,\infty)\times[R_T,\infty)$, $J(r,t)$ is non-decreasing in r and t. Moreover, when $(r,t) \in [0,R_S]\times[0,R_T]$ we have that $1 \leq J(r,t) \leq \beta/\lambda$ and when $(r,t) \in [R_S,\infty)\times[R_T,\infty)$, $J(r,t) = \beta/\lambda$. When $R_T = R_S = R > 0$, so that $S_{R_S}^{R_T} = B[0,R]$, $J(r) = J(r,r)$ is increasing and satisfies $1 \leq J(r) < \beta/\lambda$ for $r \in [0,R)$ and $J(r) = \beta/\lambda$ for $r \geq R$. For a thinned hard core process, $J_{\mathrm{inhom}}^{\mathrm{th}}(r,t) \geq 1$ for $r \leq R_S$ and $t \leq R_T$.*

Proof. Noting that $\lambda = \lambda(0,0) = \mathbb{E}[\lambda_Y(0,0;Y)] = \beta\,\mathbb{P}\left(Y \cap S_{R_S}^{R^T} = \emptyset\right) \leq \beta$, we find that $\beta/\lambda \geq 1$. Furthermore, through theorem 2 and expression (11), we obtain

$$J(r,t) = \frac{\mathbb{E}\left[\lambda_Y(0,0;Y)\mathbf{1}\left\{Y \cap S_r^t = \emptyset\right\}\right]}{\lambda(0,0)\mathbb{E}\left[\mathbf{1}\left\{Y \cap S_r^t = \emptyset\right\}\right]} = \frac{\beta}{\lambda}\frac{\mathbb{P}\left(Y \cap S_{R_S}^{R_T} = \emptyset, Y \cap S_r^t = \emptyset\right)}{\mathbb{P}\left(Y \cap S_r^t = \emptyset\right)}.$$

Hence, when both $r \geq R_S$ and $t \geq R_T$, we have that $S_{R_S}^{R_T} \subseteq S_r^t$ and consequently $J(r,t) = \beta/\lambda$. Moreover, when $r \leq R_S$ and $t \leq R_T$, so that $S_r^t \subseteq S_{R_S}^{R_T}$, expression (7) gives us $J(r,t) = 1/(1-F(r,t))$, which is increasing in both $r \in [0,R_S]$ and $t \in [0,R_T]$ and satisfies $J(r,t) \geq 1$.

Specializing to $J(r) = J(r,r)$ and $R_S = R_T = R$, when $r \leq R$, we have that $J(r) = 1/(1-F(r,r))$, which is increasing to β/λ, and when $r \geq R$, $J(r) = \beta/\lambda$.

When Y is thinned and $r \leq R_S$ and $t \leq R_T$,

$$\begin{aligned} J_{\mathrm{inhom}}^{\mathrm{th}}(r,t) =& \frac{\mathbb{E}^{!(0,0)}\left[(1-\bar{p})^{Y(S_r^t)}\right]}{\mathbb{E}\left[(1-\bar{p})^{Y(S_r^t)}\right]} \geq \frac{\mathbb{E}^{!(0,0)}\left[(1-\bar{p})^{Y\left(S_{R_S}^{R_T}\right)}\right]}{\mathbb{E}\left[(1-\bar{p})^{Y(S_r^t)}\right]} \\ =& \frac{1}{\mathbb{E}\left[(1-\bar{p})^{Y(S_r^t)}\right]} \geq 1. \end{aligned}$$

□

5.3. Log-Gaussian Cox processes

Our final example concerns spatio-temporal versions of log-Gaussian Cox processes (see, e.g., Coles & Jones (1991), Møller *et al.* (1998) or Rathbun (1996)). In words, these models are spatio-temporal Poisson processes for which the intensity functions are given by realizations of log-Gaussian random fields (Adler, 1981; Adler & Taylor, 2007).

Recall that a Gaussian random field is completely determined by its mean function $\mu(x,t)$ and its covariance function $C((x,t),(y,s)), (x,t),(y,s) \in \mathbb{R}^d \times \mathbb{R}$, and that by Bochner's theorem C must be positive definite (see, e.g., section 2.4 in Gelfand *et al.* (2010)). Now, a spatio-temporal log-Gaussian Cox process Y has random intensity function given by

$$\exp\{\mu(x,t) + Z(x,t)\}, \quad (x,t) \in \mathbb{R}^d \times \mathbb{R},$$

where $Z = \{Z(x,t)\}_{(x,t)\in\mathbb{R}^d\times\mathbb{R}}$ is a zero-mean spatio-temporal Gaussian random field. Note that the variance function of Z is given by $\sigma^2(x,t) = C((x,t),(x,t))$ and the correlation function by $r((x,t),(y,s)) = C((x,t),(y,s))/(\sigma(x,t)\sigma(y,s))$. By Daley & Vere-Jones (2003, section 6.2) or Chiu *et al.* (2013, section 5.2),

$$\frac{\rho^{(n)}((x_1,t_1),\ldots,(x_n,t_n))}{\lambda(x_1,t_1)\cdots\lambda(x_n,t_n)} = \exp\left\{\sum_{i<j} C((x_i,t_i),(x_j,t_j))\right\}$$

and the intensity function of Y is

$$\lambda(x,t) = \exp\left\{\mu(x,t) + \sigma^2(x,t)/2\right\}.$$

Therefore, if $\inf_{(x,t)} \exp\{\mu(x,t)\} > 0$ so that $\lambda(x,t)$ is bounded away from zero, under the additional condition that $C((x,t),(y,s)) = C(x-y,t-s)$, Y is IRMS. In this case, $\sigma^2(x,t) = C(0,0) = \sigma^2$ and Z is stationary. To exclude trivial cases, we shall assume that $\sigma^2 > 0$.

Before we proceed, note that we must impose conditions on r to ensure that the function $\exp\{\mu(x,t)+Z(x,t)\}$ is integrable and defines a locally finite random measure. Further details are given in the Supporting Information. Henceforth, we will assume that $\mu(x,t)$ is continuous and bounded with $\bar{\mu} = \inf_{(x,t)} \mu(x,t) > -\infty$, so that $\bar{\lambda} = \exp\left\{\bar{\mu} + \sigma^2/2\right\}$, and that r is non-negative and such that Z a.s. has continuous sample paths. Combining proposition 6.2.II in Daley & Vere-Jones (2003) with (5.35) in Chiu *et al.* (2013), under the assumptions of theorem 1, we obtain

$$J_{\text{inhom}}(r,t) = \frac{\mathbb{E}\left[\mathrm{e}^{Z(0,0)} \exp\left\{-\int_{S_r^t} \mathrm{e}^{\bar{\mu}+Z(x,s)}\,\mathrm{d}x\,\mathrm{d}s\right\}\right]}{\mathbb{E}[\mathrm{e}^{Z(0,0)}]\mathbb{E}\left[\exp\left\{-\int_{S_r^t} {}^{\bar{\mu}+Z(x,s)}\,\mathrm{d}x\,\mathrm{d}s\right\}\right]}$$

upon noting that the Palm distribution of the driving random measure of our log-Gaussian Cox process Y is e^Z-weighted. Note here that the Papangelou conditional intensity of Y exists and is given by $\lambda(x,t;Y) = \mathbb{E}[\exp\{\mu(x,t) + Z(x,t)\}|Y]$ (see, e.g. Møller & Waagepetersen (2007)).

Lemma 3. *For a log-Gaussian Cox process, when the aforementioned conditions are imposed on* μ *and* C, $J_{\text{inhom}}(r,t) \leq 1$ *for all* $r,t \geq 0$.

Proof. First, observe that $J_{\text{inhom}}(r,t) \leq 1$ is equivalent to

$$\mathrm{Cov}\left(\mathrm{e}^{Z(0,0)}, \exp\left\{-\mathrm{e}^{\bar{\mu}} \int_{S_r^t} \mathrm{e}^{Z(x,s)}\,\mathrm{d}x\,\mathrm{d}s\right\}\right) \leq 0.$$

Further, note that by the a.s. sample path continuity of Z,

$$\exp\left\{-\mathrm{e}^{\bar{\mu}} \int_{S_r^t} \mathrm{e}^{Z(x,s)}\,\mathrm{d}x\,\mathrm{d}s\right\} \stackrel{\text{a.s.}}{=} \lim_{n\to\infty} \exp\left\{-\mathrm{e}^{\bar{\mu}} \sum_{(x_i,s_i)\in S(n)} c_{i,n}\mathrm{e}^{Z(x_i,s_i)}\right\},$$

where $S(n) \subseteq S_r^t, n \geq 1$, are Riemann partitions. Because Z has positive correlation function, Pitt's theorem (Pitt, 1982) tells us that Z is an associated family of random variables. Hereby, $\mathrm{Cov}\left(e^{Z(0,0)}, \exp\left\{-e^{\bar{\mu}} \sum_{(x_i,s_i)\in S(n)} c_{i,n} e^{Z(x_i,s_i)}\right\}\right) \leq 0$ for any $n \geq 1$, and the result follows from taking the limit in the last covariance and applying dominated convergence. □

6. Estimation

Assume that we observe an IRMS STPP Y within some compact spatio-temporal region $W_S \times W_T \subseteq \mathbb{R}^d \times \mathbb{R}$ and obtain the realization $\{(x_i, t_i)\}_{i=1}^m, m = Y(W_S \times W_T)$. The goal of this section is to derive estimators for $G_{\text{inhom}}(r,t)$, $F_{\text{inhom}}(r,t)$ and $J_{\text{inhom}}(r,t)$. In order to deal with possible edge effects, we will apply a minus sampling scheme (Chiu *et al.*, 2013; Cronie & Särkkä, 2011). For clarity of exposition, we assume that the intensity function is known.

Denote the boundaries of W_S and W_T by ∂W_S and ∂W_T, respectively. Further, write $W_S^{\ominus r} = \{x \in W_S : d_{\mathbb{R}^d}(x, \partial W_S) \geq r\} = \{x \in W_S : x + B_{\mathbb{R}^d}[0,r] \subseteq W_S\}$ for the eroded spatial domain, and similarly, let $W_T^{\ominus t} = \{s \in W_T : d_{\mathbb{R}}(s, \partial W_T) \geq t\}$. For given $r, t \geq 0$, we define an estimator of $1 - G_{\text{inhom}}(r,t)$ by

$$\frac{1}{\left|Y \cap \left(W_S^{\ominus r} \times W_T^{\ominus t}\right)\right|} \sum_{y \in Y \cap \left(W_S^{\ominus r} \times W_T^{\ominus t}\right)} \left[\prod_{(x,s)\in(Y\setminus\{y\})\cap(y+S_r^t)} \left(1 - \frac{\bar{\lambda}}{\lambda(x,s)}\right)\right] \tag{12}$$

and, given a finite point grid $L \subseteq W_S \times W_T$, we estimate $1 - F_{\text{inhom}}(r,t)$ by

$$\frac{1}{\left|L \cap \left(W_S^{\ominus r} \times W_T^{\ominus t}\right)\right|} \sum_{l \in L \cap \left(W_S^{\ominus r} \times W_T^{\ominus t}\right)} \left[\prod_{(x,s)\in Y\cap(l+S_r^t)} \left(1 - \frac{\bar{\lambda}}{\lambda(x,s)}\right)\right]. \tag{13}$$

The ratio of (12) and (13) is an estimator of $J_{\text{inhom}}(r,t)$ (cf. theorem 1).

Theorem 3. *Under the conditions of theorem 1, the estimator (13) is unbiased and (12) is ratio unbiased.*

Proof. We start with (12) and note that $\mathbb{E}\left[Y\left(W_S^{\ominus r} \times W_T^{\ominus t}\right)\right] = \Lambda\left(W_S^{\ominus r} \times W_T^{\ominus t}\right)$. By the reduced Campbell–Mecke formula (4),

$$\mathbb{E}\left[\sum_{y \in Y \cap \left(W_S^{\ominus r} \times W_T^{\ominus t}\right)} \prod_{(x,s)\in(Y\setminus\{y\})\cap(y+S_r^t)} \left(1 - \frac{\bar{\lambda}}{\lambda(x,s)}\right)\right]$$

$$= \int_{W_S^{\ominus r} \times W_T^{\ominus t}} \mathbb{E}^{!y}\left[\prod_{(x,s)\in Y} \left(1 - \frac{\bar{\lambda}}{\lambda(x,s)} \mathbf{1}\left\{(x,s) \in y + S_r^t\right\}\right)\right] \lambda(y)\, \mathrm{d}y.$$

By (9), the expectation is equal to $G^{!0}\left(1 - u_{r,t}^0\right)$, from which the claimed ratio-unbiasedness follows.

Turning to (13), unbiasedness follows from the assumed translation invariance of the ξ_n and equation (6) under the conditions of theorem 1. □

Estimators for K^*_{inhom} follow immediately from the definition and are discussed in Gabriel & Diggle (2009).

In the Supporting Information we use the inhomogeneous J-functions and K-functions to quantify the interactions in realizations of each of the three models discussed in section 5. We work mostly in **R**. For the J-function, we exploit functions in the package `spatstat` (Baddeley & Turner, 2005); an estimator for K^*_{inhom} has been implemented in `stpp` (Gabriel *et al.*, 2013). To simulate log-Gaussian Cox processes, we use the package `RandomFields` by Schlather *et al.* (2013). Realizations of spatio-temporal hard core processes can be obtained using the `C++` library `MPPLIB` of Steenbeek *et al.* (2002). Because the intensity function is either known or known up to a constant (for the thinned hard core process), as (12)–(13) are defined in terms of the ratio $\bar{\lambda}/\lambda(x,s)$, there is no need to plug in intensity function estimators.

In practice, the intensity function $\lambda(x,s)$ and its infimum are unknown and must be estimated. Indeed, Gabriel & Diggle (2009) as well as Møller & Ghorbani (2012) consider a combination of parametric models and kernel estimators for $\lambda(x,s)$. They stress, however, that care has to be taken when $\hat{\lambda}(x,s)$ is close to zero.

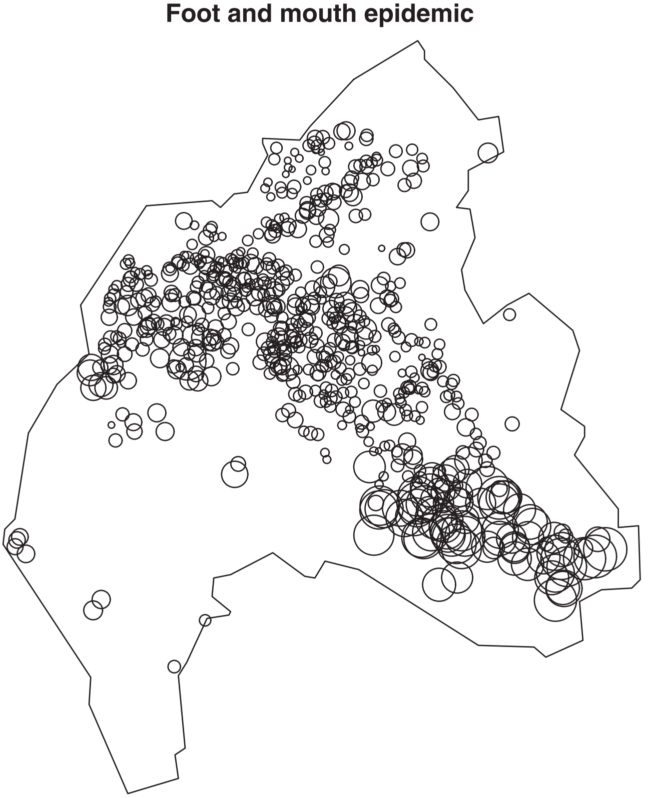

Fig. 1. Incidences of foot and mouth disease in northern Cumbria. The locations of outbreak are indicated by the centre of a disc whose radius is proportional to the date of the outbreak in days from 1 February 2001.

7. Application

Figure 1 plots incidences of foot and mouth disease in the northern part of the county of Cumbria in the west of England bordering Scotland. Because the original data contain confidential information, we use the publicly available version in the **R** package `stpp` (Gabriel *et al.*, 2013). The locations of outbreak are indicated by the centre of a disc whose radius is proportional to the date of the outbreak in days from 1 February 2001. For the spatial window W_S, we take the territory outlined in Figure 1; as data were collected for 200 days, we set $W_T = [0, 200]$.

We follow Møller & Ghorbani (2012) and assume that the unknown intensity function λ is separable. The spatial intensity is estimated by a mass preserving Gaussian kernel estimator (Lieshout, 2012) with standard deviation (bandwidth) 3.68 km, the temporal intensity by a log-transform re-transform scheme (Markovich, 2007) using a Gaussian kernel with standard deviation 0.05 days. For further details, see Møller & Ghorbani (2012). The bound $\bar{\lambda}$ was estimated by the minimal value of $\hat{\lambda}$ among the data points.

Figure 2 shows $\widehat{J_{\text{inhom}}(r)}$ for the Cumbria data (black line) together with envelopes obtained from 99 independent samples from a spatio-temporal Poisson process with intensity $\hat{\lambda}$. Note that the mean of the latter (stippled line) is close to one and that the variation increases with r. There is clear evidence of clustering, confirming findings from Gabriel *et al.* (2013) and Møller & Ghorbani (2012) based on $\widehat{K^*_{\text{inhom}}}$.

It should be stressed that we do not formally test hypotheses here but rather look for indications of clustering. Provided that one formally would like to perform such tests, it might be advised to consider the testing schemes considered by Grabarnik *et al.* (2011) and Myllymäki *et al.* (2013). However, these tests require very large numbers of simulated three-dimensional

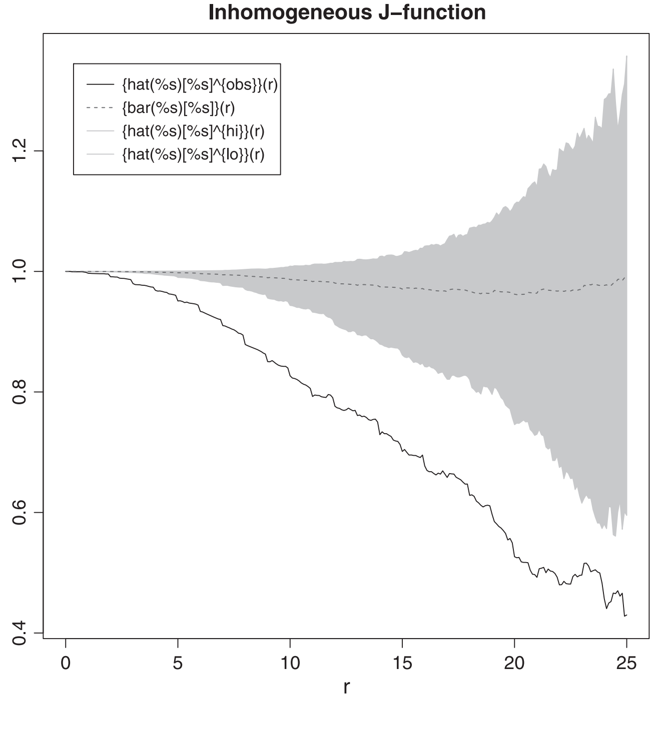

Fig. 2. Estimated $J_{\text{inhom}}(r)$-values for the Cumbria data (black line) together with envelopes obtained from 99 independent samples from a spatio-temporal Poisson process with the same (estimated) intensity.

point patterns to be used in the envelopes. As a consequence, we strongly believe that such computations would be virtually impossible to perform here.

8. Discussion

In this paper, we followed the suggestion of Lieshout (2011) to generalize the nearest neighbour distance distribution, empty space and J-functions for inhomogeneous point patterns to the spatio-temporal domain by equipping $\mathbb{R}^d \times \mathbb{R}$ with the supremum distance and applying scaling under an appropriate assumption of intensity-reweighted stationarity. We expressed the new statistics in terms of fundamental point process characteristics including product densities, the generating functional and the Papangelou conditional intensity. We computed the inhomogeneous J-function for three important families of STPPs and discussed the relation to the inhomogeneous spatio-temporal K-function. Finally, we derived non-parametric estimators and illustrated their performance by means of a data set on the 2001 foot and mouth epidemic in the UK.

In practice, the intensity function tends to be unknown and must be estimated. This is not a problem when there are independent replicates. Otherwise, pragmatic model assumptions must be made. For example, in section 7, we imposed separability. When prior information about the data is available, a parametric model can also be used.

Throughout, an STPP Y was understood as a point process on a product space. However, when one of the dimensions is of prime importance, it would be natural to treat Y as a marked point process. Thus, our work in progress aims at defining versions of the G-functions, J-functions and K-functions for inhomogeneous Cronie & Lieshout (2014) marked point processes. The focus will be on real-valued or discrete marks, but it is important to bear in mind that more complicated cases, for example the function spaces of Cronie & Mateu (2014), can be handled as well.

Acknowledgements

The authors are grateful to two anonymous referees and to the editor for useful comments, and to Guido Legemaate, Jesper Møller and Alfred Stein for interesting discussions. The authors thank Mohammad Ghorbani for code relating to (Møller & Ghorbani, 2012); Martin Schlather for access to and help with an updated version of his **R** package `RandomFields`; and Peter Diggle, Edith Gabriel and Barry Rowlingson for the data set on the UK foot and mouth epidemic in their **R** package `stpp`. This research was supported by the Netherlands Organisation for Scientific Research NWO (613.000.809).

References

Adler, R. J. (1981). *The geometry of random fields*, Wiley, Chichester.

Adler, R. J. & Taylor, J. E. (2007). *Random fields and geometry*, Springer, New York.

Baddeley, A. J., Møller, J. & Waagepetersen, R. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statist. Neerlandica* **54**, 329–350.

Baddeley, A. & Turner, R. (2005). Spatstat: an R package for analyzing spatial point patterns. *J. Stat. Softw.* **12**, 1–42.

Bedford, T. & Van den Berg, J. (1997). A remark on the Van Lieshout and Baddeley J-function for point processes. *Adv. in Appl. Probab.* **29**, 19–25.

Chiu, S. N., Stoyan, D., Kendall, W. S. & Mecke, J. (2013). *Stochastic geometry and its applications*, (3rd edn)., Wiley, Chichester.

Coles, P. & Jones, B. (1991). A lognormal model for the cosmological mass distribution. *Mon. Not. R. Astron. Soc.* **248**, 1–13.

Cronie, O. & Lieshout, M. N. M. van. (2014). Summary statistics for inhomogeneous marked point processes. Available at: http://arxiv.org/abs/1407.3089 [Accessed on 26 September 2014].

Cronie, O. & Särkkä, A. (2011). Some edge correction methods for marked spatio-temporal point process models. *Comput. Statist. Data Anal.* **55**, 2209–2220.

Cronie, O. & Mateu, J. (2014). *Spatio-temporal càdlàg functional marked point processes: unifying spatio-temporal frameworks*. Available at: http://arxiv.org/abs/1403.2363 [Accessed on 26 September 2014].

Daley, D. J. & Vere-Jones, D. (2003). *An introduction to the theory of point processes: volume I: elementary theory and methods*, (2nd edn)., Springer, New York.

Daley, D. J. & Vere-Jones, D. (2008). *An introduction to the theory of point processes: volume II: general theory and structure*, (2nd edn)., Springer, New York.

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

Gabriel, E., Rowlingson, B. S. & Diggle, P. J. (2013). STPP: an R package for plotting, simulating and analyzing spatio-temporal point patterns. *J. Stat. Softw.* **53**, 1–29.

Gelfand, A. E., Diggle, P. J., Fuentes, M. & Guttorp, P. (2010). *Handbook of spatial statistics*, Taylor & Francis, Boca Raton.

Grabarnik, P., Myllymäki, M. & Stoyan, D. (2011). Correct testing of mark independence for marked point patterns. *Ecol. Model.* **222**, 3888–3894.

Halmos, P. R. (1974). *Measure theory*, Springer, New York.

Illian, J., Penttinen, A., Stoyan, H. & Stoyan, D. (2008). *Statistical analysis and modelling of spatial point patterns*, Wiley, Chichester.

Lieshout, M. N. M. van. (2000). *Markov point processes and their applications*, Imperial College Press/World Scientific, London/Singapore.

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

Lieshout, M. N. M. van. (2012). On estimation of the intensity function of a point process. *Methodol. Comput. Appl. Probab.* **14**, 567–578.

Lieshout, M. N. M. van & Baddeley, A. J. (1996). A nonparametric measure of spatial interaction in point patterns. *Statist. Neerlandica* **50**, 344–361.

Markovich, N. (2007). *Nonparametric analysis of univariate heavy-tailed data: research and practice*, Wiley, Chichester.

Møller, J. & Díaz–Avalos, C. (2010). Structured spatio-temporal shot-noise Cox point process models, with a view to modelling forest fires. *Scand. J. Statist.* **37**, 2–25.

Møller, J. & Ghorbani, M. (2012). Aspects of second-order analysis of structured inhomogeneous spatio-temporal point processes. *Statist. Neerlandica* **66**, 472–491.

Møller, J., Syversveen, A. R. & Waagepetersen, R. P. (1998). Log Gaussian Cox processes. *Scand. J. Statist.* **25**, 451–482.

Møller, J. & Waagepetersen, R. P. (2007). Modern statistics for spatial point processes (with discussion). *Scand. J. Statist.* **34**, 643–711.

Myllymäki, M., Mrkvicka, T., Seijo, H. & Grabarnik, P. (2013). *Global envelope tests for spatial processes*. Available at: http://arxiv.org/pdf/1307.0239.pdf [Accessed on 26 September 2014].

Ogata, Y. (1998). Space–time point-process models for earthquake occurrences. *Ann. Inst. Statist. Math.* **50**, 379–402.

Pitt, L. D. (1982). Positively correlated normal variables are associated. *Ann. Probab.* **10**, 496–499.

Rathbun, S. L. (1996). Estimation of Poisson intensity using partially observed concomitant variables. *Biometrics* **52**, 226–242.

Schlather, M., Menck, P., Singleton, R. & Pfaff, B. (2013). *RandomFields: simulation and analysis of random fields*. Available at: http://cran.r-project.org/web/packages/RandomFields/index.html [Accessed on 26 September 2014].

Schneider, R. & Weil, W. (2008). *Stochastic and integral geometry*, Springer, Berlin.

Steenbeek, A. G., Lieshout, M. N. M. van, Stoica, R. S. with contributions from, Gregori, P. & Berthelsen, K. K. (2002). *MPPBLIB, a C++ library for marked point processes*, CWI, Amsterdam.

White, S. D. M. (1979). The hierarchy of correlation functions and its relation to other measures of galaxy clustering. *Mon. Not. R. Astron. Soc.* **186**, 145–154.

Received November 2013, in final form September 2014

M.N.M. van Lieshout, CWI, Amsterdam, The Netherlands.
E-mail: M.N.M.van.Lieshout@cwi.nl

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site including:

Appendix S1: Proofs of the theorems in section 4;
Appendix S2: Sample path continuity of Gaussian random fields;
Appendix S3: Covariance models;
Appendix S4: Numerical evaluations.