Discussion of the Paper "Marked Spatial Point Processes: Current State and Extensions to Point Processes on Linear Networks"

Ottmar CRONIE, Julia JANSSON, and Konstantinos KONSTANTINOU

The authors of the paper under discussion have done an excellent job in providing a summary of the current state of a large part of the field, as well as proposing extensions to marked linear network point processes. Both are welcomed additions to the literature, and we believe that the paper will be a valuable resource for the spatial statistics community.

The title of the paper might give the impression that the paper is a review of "all of (statistics for)" marked point processes, which in itself would have been quite an endeavour to complete, given the long history of the field. Fortunately, the authors restrict themselves to reviewing summary statistics for point processes, which is more manageable, given the page limitation. For the reader who anticipates a full review of statistics for marked point processes, we would, however, like to highlight a few important parts of the field. It should be noted that this list is far from complete in terms of additional topics.

- Marked temporal point processes, e.g. Hawkes and ETAS processes, which are expressed through temporal conditional intensity functions, have a long history and have the advantage that the likelihood function is known in closed form (Daley and Vere-Jones 2003, 2008). This, in turn, leads to a statistical analysis much in line with classical statistics. A fairly recent review on the topic can be found in Reinhart (2018).
- There is a long trajectory of papers dealing with the Growth-Interaction process of Renshaw and Särkkä (Särkkä and Renshaw 2006; Renshaw et al. 2009; Lavancier and Le Guével 2021; Redenbach and Särkkä 2013; Cronie et al. 2013; Cronie and Särkkä 2011; Renshaw and Särkkä 2001; Cronie 2012). This is a temporally dynamic spatial model which allows for the generation of point processes with dynamic real-valued marks, with various types of interactions structures.
- There is a vast literature on marked point process models and their statistical inference, e.g. van Lieshout (2000), Goulard et al. (1996), Grabarnik and Särkkä (2009), Hög-

This is a commentary to the article https://doi.org/10.1007/s13253-024-00605-1.

O. Cronie (✉) · J. Jansson · K. Konstantinou, Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, Göteborg, Sweden (E-mail: *ottmar@chalmers.se*).

Journal of Agricultural, Biological, and Environmental Statistics, Volume 29, Number 2, Pages 379–388
https://doi.org/10.1007/s13253-024-00606-0

mander and Särkkä (1999), Billiot et al. (2008), Coeurjolly et al. (2012), Coeurjolly and Lavancier (2013), to mention a few Gibbs-process-related works. Clearly, this list is far from complete, but it offers a point of departure for interested readers.

- Estimators of summary statistics for marked point processes serve different purposes. Perhaps most notably, they may be used to test for independence between marks, points and covariates; see e.g. Lotwick and Silverman (1982), Cronie and van Lieshout (2016), Mrkvička et al. (2021), Dvořák et al. (2022).

The remainder of this discussion will be devoted to clarifying a few aspects which people may find hard to grasp and to discussing an appropriate setup for summary statistics for general marked point processes on linear networks.

1. MARK SPACES

Given a marked point process $X = \{(x_i, m(x_i))\}_{i=1}^N$ with spatial domain $\mathcal{S}$ (in the paper under discussion, either a Euclidean domain or a linear network) and mark space $\mathcal{M}$, the general theory for marked point processes, which has been detailed in e.g. the monographs of van Lieshout (2000), Daley and Vere-Jones (2003), Daley and Vere-Jones (2008), covers point processes where both the spatial domain and the mark space are arbitrary complete separable metric (csm) spaces. Typical examples of such spaces include the following:

- Using the Euclidean metric $d_{\mathcal{M}}(m_1, m_2) = \|m_1 - m_2\|$, $m_1, m_2 \in \mathcal{M}$, for either $\mathcal{M} = \{1, \dots, k\}$, which results in a multi-type point process, or $\mathcal{M} = \mathbb{R}^d$, which corresponds to a point process with real-valued marks, we obtain a csm space $\mathcal{M}$. Note that one may identify a multi-type point process X with the multivariate point process $(X_1, \dots, X_k)$, where $X_j = \{x_i : (x_i, m(x_i)) \in X, m(x_i) = j\}$, i.e. a vector of k unmarked point processes.
- More abstract csm mark spaces are also possible. A first example is when $\mathcal{M}$ is given by a function space $\mathcal{F}$, equipped with a suitable metric $d_{\mathcal{M}}$; see Ghorbani et al. (2021) for details. The simplest example is when $\mathcal{F}$ is given by all continuous functions $f : \mathbb{R}^k \to \mathbb{R}^{k'}$, where either $k' = 1$ or $k' = k$, and $d_{\mathcal{M}}(m_1, m_2) = \sup_{t_1, t_2 \in \mathcal{R}^k} |m_1(t_1) - m_2(t_2)|$. Other, more advanced, examples include L_p-spaces and Skorohod spaces.
- In stochastic geometry, it is common to encounter marks which are given by subsets of $\mathbb{R}^d$. This may be achieved by letting the mark space $\mathcal{M}$ be given by the collection of closed subsets, equipped with the metric generating the so-called Fell topology (Molchanov 2005). Examples here include germ-grain/Boolean models.
- One can straightforwardly build quite elaborate mark spaces $\mathcal{M} = \mathcal{M}_1 \times \cdots \times \mathcal{M}_n$, given by products of different csm spaces (Daley and Vere-Jones 2003, 2008; Ghorbani et al. 2021).

Before we proceed to discussing different summary statistics in detail, we believe that it is important to have an understanding of what a pair of (Borel) sets $C, D \subseteq \mathcal{M}$ may look like, for the first four examples of $\mathcal{M}$ above:

- When $\mathcal{M} = \{1, \dots, k\}$, we may for instance consider $C = \{i\}$ for some $i \in \{1, \dots, k\}$, and, say, $D = \{j\}$, $j \neq i$, or $D = \{1, \dots, k\} \setminus \{i\}$.
- When $\mathcal{M} = \mathbb{R}$, we may for instance be interested in comparing points with marks in the intervals $C = [0, 10]$ and $D = (10, 20]$.
- If the mark space $\mathcal{M} = \mathcal{C}[0, T]$, $T > 0$, is the collection of continuous functions $f : [0, T] \to \mathbb{R}$, we may be interested in comparing points with marks in $C = \{f \in \mathcal{M} : f(t) < 10 \text{ for all } t \in \mathbb{R}\}$ and $D = \{f \in \mathcal{M} : f(t) > 5 \text{ for } t \in [T/2, T]\}$.
- Let $\mathcal{M}$ be given by the collection of closed sets in $\mathbb{R}^d$. Examples of mark sets could be $C = \{B \in \mathcal{M} : B \cap W_1 \neq \emptyset\}$ and $D = \{B \in \mathcal{M} : B \cap W_2 = \emptyset\}$ for some fixed $W_1, W_2 \subseteq \mathbb{R}^d$.

The mathematical tools used to define the summary statistics presented in the paper under discussion, e.g. product density/intensity functions for marked point processes, are all valid for general csm mark spaces. Hence, from a theoretical/methodological point of view they can be employed in all kinds of settings. It is mainly in the implementation context that details about the underlying spaces become relevant.

2. SUMMARY STATISTICS

Roughly speaking, most summary statistics fall into two different categories. On the one hand, there are those that filter the point process into mark groups, based on fixed marks sets, and then reflect spatial dependencies between the points of these groups. On the other hand, there are the mark-correlation type statistics, which condition on the presence of two points (a given distance apart) and reflect second moment properties of their associated marks.

2.1. PRODUCT DENSITY AND INTENSITY FUNCTIONS

Given a marked point process $X = \{(x_i, m(x_i))\}_{i=1}^N$, where both the spatial domain $\mathcal{S}$ and the mark space $\mathcal{M}$ are csm spaces, most of the statistical tools available can be described through its product densities

$$\lambda^{(n)}((x_1, m_1), \dots, (x_n, m_n)) = f_{\mathcal{M}}^{(n)}(m_1, \dots, m_n | x_1, \dots, x_n) \lambda_{\mathcal{S}}^{(n)}(x_1, \dots, x_n), \quad n \geq 1,$$

which govern the probabilities that X has a non-empty intersection with infinitesimal neighbourhoods $d(x_1, m_1) \times \cdots \times d(x_n, m_n)$. Here, $\lambda_{\mathcal{S}}^{(n)}$ is the nth-order product density of the ground/unmarked point process $X_{\mathcal{S}} = \{x_i\}_{i=1}^N$ and $f_{\mathcal{M}}^{(n)}$ is a conditional density for n marks, given that the associated ground points have the spatial locations $x_1, \dots, x_n$. When $n = 1$, the product densities are called intensity functions, and when $\lambda_{\mathcal{S}}^{(1)}(x)$ is (non)constant over $\mathcal{S} \times \mathcal{M}$ we say that X is (in)homogeneous; if X is stationary, it is automatically homogeneous. Moreover, if the marks are independent conditionally on the ground process, then $f_{\mathcal{M}}^{(n)}(m_1, \dots, m_n | x_1, \dots, x_n) = f_{\mathcal{M}}^{(1)}(m_1 | x_1) \cdots f_{\mathcal{M}}^{(1)}(m_1 | x_1)$.

It is important to understand that product densities are defined as densities of factorial moment measures, with respect to reference measures. Consequently, one has to make choices for reference measures on $\mathcal{S}$ and $\mathcal{M}$. The natural choice for the former is Lebesgue measure when $\mathcal{S} = \mathbb{R}^n$ and arc-length integration when $\mathcal{S}$ is a linear network. For the reference measure $\nu_{\mathcal{M}}$ on the mark space there are some canonical choices: for Euclidean mark spaces we use Lebesgue measure and for a multi-type point process we let $\nu_{\mathcal{M}}$ assign the value 1 to each element of $\mathcal{M} = \{1, \dots, k\}$, i.e. it is the counting measure, whereby integrals become sums. Hereby, in the case of a multi-type point process, $f_{\mathcal{M}}^{(n)}(\cdot|x_1, \dots, x_n)$ is a multivariate probability mass function on $(1, \dots, k)^n$, while for real-valued marks it is a multivariate density function on $(\mathbb{R}^d)^n$. Clearly, when $\mathcal{M}$ is, say, a function space or a collection of closed Euclidean subsets, the notion of a density function on $\mathcal{M}^n$, which is mathematically sound in itself, becomes very abstract and hard to grasp, not to mention practically challenging (Ghorbani et al. 2021).

2.2. van Lieshout (2006): Mark Filtering under Stationarity

The idea of mark filtering with respect to mark sets belonging to a general csm space $\mathcal{M}$ was, to the best of our knowledge, introduced by van Lieshout (2006). For marked point processes with spatial locations in $\mathcal{S} = \mathbb{R}^n$, van Lieshout considered a general marked K-function which reflects the expected number of further points within radius r of a typical point x_i with mark $m(x_i) \in C$, for a fixed mark set C. Indeed, this may be seen as a generalisation of the classical i-to-any K-function for stationary multivariate/multi-type point processes. More noteworthy, perhaps, is that van Lieshout's K-function is obtained as a second-order approximation of van Lieshout's marked J-function, a higher-order statistic, which constitutes a combination of van Lieshout's marked versions of the nearest-neighbour distance distribution function and the empty space function.

2.3. Notions of Marked Intensity Reweighted Stationarity

In the inhomogeneous setting, instead of classical stationarity, some alternative form of transformation invariance of the distributional properties of X is required. Following Cronie and van Lieshout (2016) and Ghorbani et al. (2021), under the assumption that $\mathcal{S} = \mathbb{R}^n$ and $\mathcal{M}$ is a csm space, the natural way forward is to impose that the intensity function is bounded away from 0 and that

$$
\begin{aligned}
g^{(n)}((x_1, m_1), \dots, ((x_n, m_n)) &= \frac{\lambda^{(n)}((x_1, m_1), \dots, (x_n, m_n))}{\lambda^{(1)}(x_1, m_1) \cdots \lambda^{(1)}(x_n, m_n)} \\
&= \frac{\lambda_{\mathcal{S}}^{(n)}(x_1, \dots, x_n)}{\lambda_{\mathcal{S}}^{(1)}(x_1) \cdots \lambda_{\mathcal{S}}^{(1)}(x_n)} \frac{f_{\mathcal{M}}^{(n)}(m_1, \dots, m_n|x_1, \dots, x_n)}{f_{\mathcal{M}}^{(1)}(m_1|x_1) \cdots f_{\mathcal{M}}^{(1)}(m_n|x_n)} \\
&= g_{\mathcal{S}}^{(n)}(x_1, \dots, x_n) \gamma_{\mathcal{M}}^{(n)}(m_1, \dots, m_n|x_1, \dots, x_n)
\end{aligned}
$$

is translation invariant in its argument $(x_1, \dots, x_n)$ for sufficiently many orders n; it holds automatically for $n = 1$. If this is true for all orders $n \leq k$, we speak of kth-order (marked) intensity reweighted stationarity (k-IRS) and when this is true for any $n \geq 1$, we call X (marked) intensity reweighted moment stationary (IRMS). Note that the translation invariance pertains to both $g_{\mathcal{S}}^{(n)}$ and $\gamma_{\mathcal{M}}^{(n)}$ and further that $\gamma_{\mathcal{M}}^{(n)}$ vanishes under independent marking.

We believe that it is important to emphasise these concepts, as they are both intricate and central when dealing with summary statistics for marked point processes.

2.4. Cronie and van Lieshout (2016): Mark Filtering under Inhomogeneity

Building upon the ideas of van Lieshout (2011) and Cronie and van Lieshout (2016) extended the summary statistics of van Lieshout (2006), under the assumption that X is IRMS. Here, in the case of a general csm mark space $\mathcal{M}$, the aim is to quantify spatial dependencies between points that belong to different mark subsets $C, D \subseteq \mathcal{M}$, having adjusted for the varying intensity. More specifically, under sufficient regularity of $g^{(n)}, n \geq 1$, their marked inhomogeneous J-function is given by $J_{\text{inhom}}^{CD}(r) = (1 - G_{\text{inhom}}^{CD}(r))/(1 - F_{\text{inhom}}^{D}(r))$, where, given $\bar{\lambda}_D = \inf_{z \in \mathbb{R}^d, m \in D} \lambda^{(1)}(z, m)$ and

$$f_D^u(r, X) = \prod_{(x,m) \in X} \left(1 - \frac{\bar{\lambda}_D \mathbf{1}\{\|u - x\| \leq r\} \mathbf{1}\{m \in D\}}{\lambda^{(1)}(x, m)}\right), \quad u \in \mathcal{S} = \mathbb{R}^n,$$

the marked inhomogeneous versions of the nearest-neighbour distance distribution function, $G_{\text{inhom}}^{CD}(r) = 1 - \nu_{\mathcal{M}}(C)^{-1} \int_C \mathbb{E}^!_{(u,m')}[f_D^u(r, X)] \nu_{\mathcal{M}}(dm')$, and the empty space function, $F_{\text{inhom}}^{D}(r) = 1 - \mathbb{E}^!_{(u,m')}[f_D^u(r, X)]$, each takes the same value for (almost) any u. Here, $\mathbb{E}^!_{(u,m')}[\cdot]$ is the expectation corresponding to the distribution of the reduced Palm process $X^!_{(u,m')}$, $(u, m') \in \mathbb{R}^n \times \mathcal{M}$, which we interpret as X conditioned on having a point in (u, m') which we remove upon realisation. The former reflects the distribution that there are no points with marks in D within distance r from another point with mark in C, whose existence we (Palm) condition on, when we have scaled away the effect of the varying intensity function. $F_{\text{inhom}}^{D}(r)$ essentially reflects the same, but here the point with mark in C is replaced by an arbitrary location $u \in \mathbb{R}^n$. By Slivnyak's theorem, the two coincide for a (marked) Poisson process on $\mathbb{R}^n \times \mathcal{M}$.

So why do we need to impose IRMS? The answer is that the components of $J_{\text{inhom}}^{CD}(r)$ are expressed as evaluations of generating functionals (Daley and Vere-Jones 2008), which may be expanded in terms of the functions $g^{(n)}$, $n \geq 1$. The imposed translation invariance of these functions, i.e. IRMS, ensures, among other things, that the ratio of the expansions for $1 - G_{\text{inhom}}^{CD}(r)$ and $1 - F_{\text{inhom}}^{D}(r)$ gives us the expansion for $J_{\text{inhom}}^{CD}(r)$, and that the (Palm) conditioning in $G_{\text{inhom}}^{CD}(r)$ is location independent. The latter, in turn, allows us to properly construct an estimator for $G_{\text{inhom}}^{CD}(r)$ as a sum which runs over the points of the point process in question; we need this since, in practice, we do not have repeated sampling of the point process. The same holds true for e.g. second-order IRS in the case of the (spatiotemporal) inhomogeneous K-function or IRMS in the case of (spatiotemporal) higher-order summary statistics for unmarked processes (Baddeley et al. 2000; Gabriel and Diggle 2009; van Lieshout 2011; Møller and Ghorbani 2012; Cronie and van Lieshout 2015).

2.5. GHORBANI ET AL. (2021) AND D'ANGELO ET AL. (2023): GENERAL MARK-WEIGHTING UNDER INHOMOGENEITY

For a k-IRS marked point process with $\mathcal{S} = \mathbb{R}^n$ and csm mark space $\mathcal{M}$, Ghorbani et al. (2021) proposed the, arguably, most general form of a finite-order summary statistic, a mark-weighted kth-order summary statistic. We here illustrate it when $k = 2$ (i.e. we focus on second-order interactions):

$$\begin{aligned}
\mathcal{K}_t(A) &= \frac{1}{|W|}\mathbb{E}\left[\sum_{(x,m)\in X\cap W\times\mathcal{M}}\ \sum_{(x',m')\in X\setminus\{(x,m)\}} \frac{\mathbf{1}\{x'-x\in A\}t(m,m')}{\lambda^{(1)}(x,m)\lambda^{(1)}(x',m')}\right] \\
&= \int_{\mathcal{M}}\int_{A\times\mathcal{M}} t(m,m')g^{(2)}((0,m),(x',m'))dx'\nu_{\mathcal{M}}(dm')\nu_{\mathcal{M}}(dm) \\
&= \int_{A}\int_{\mathcal{M}^2} t(m,m')\gamma^{(2)}_{\mathcal{M}}(m,m'|0,x')\nu_{\mathcal{M}}(dm')\nu_{\mathcal{M}}(dm)g^{(2)}_{\mathcal{S}}(0,x')dx' \\
&= \int_{\mathcal{M}} \mathbb{E}^{!}_{(x,m)}\left[\sum_{(x',m')\in X\cap(x+A)} \frac{t(m,m')}{\lambda^{(1)}(x',m')}\right]\nu_{\mathcal{M}}(dm).
\end{aligned}$$

The mapping $t : \mathcal{M}^2 \to \mathbb{R}$ is a test function which governs by what quantity $t(m,m')$ we weight the contribution of a pair (x,m), (x',m') to the summary statistic. Moreover, the indicator function $\mathbf{1}\{x - x' \in A\}$ ensures that for each (x,m) we only consider its A-neighbours, i.e. those $(x',m') \in X\setminus\{(x,m)\}$, satisfying that $x' \in (x + A)$. We here see that, because of k-IRS, the definition is independent of the choice of $W \subseteq \mathcal{S} = \mathbb{R}^n$, $|W| > 0$. In addition, we have that $\mathcal{K}_t(A)$ does not change with the reduced Palm conditioning in $x \in \mathbb{R}^n$. Note that nothing here restricts us to a particular kind of csm mark space. Regarding further generalisations, one could, of course, also look into replacing $\mathbf{1}\{x - x' \in A\}t(m,m')$ by some general weight function $t : (\mathbb{R} \times \mathcal{M})^2 \to \mathbb{R}$.

Turning to special cases of $\mathcal{K}_t(A)$, when we let $t(m,m') = \mathbf{1}\{m \in C, m' \in D\}$ for two mark sets $C, D \subseteq \mathcal{M}$, we retrieve the marked second-order reduced moment measure $\mathcal{K}^{CD}(A)$, $A \subseteq \mathbb{R}^n$, of Iftimi et al. (2019). By further letting $A = b(0,r)$ be the closed origin-centred r-ball in $\mathbb{R}^n$, we retrieve the marked inhomogeneous K-function of Cronie and van Lieshout (2016), while e.g. elliptical origin-centred sets A allow us to analyse anisotropic interactions between points with marks in C and D (Iftimi et al. 2019). In the multi-type setting, by letting $C = \{i\}$ and $D = \{j\}$, or $D = \mathcal{M} \setminus \{i\}$, we obtain inhomogeneous variants of classical cross summary statistics, as indicated in the review part of the paper under discussion.

Now, consider the setting where the marginal distributions of the marks are all the same, which e.g. is the case when the marks are iid, when X is stationary, or when the marking function $m(x) \in \mathcal{M}$, $x \in \mathbb{R}^n$, is a stationary random field. If now, in addition, this common marginal distribution is used as reference measure $\nu_{\mathcal{M}}$ for the mark space $\mathcal{M}$, then $\lambda^{(1)}(x,m) = \lambda^{(1)}_{\mathcal{S}}(x)$. This assumption is particularly convenient in statistical settings when $\mathcal{M}$ is abstract, e.g. a function space or a space of closed sets, since it circumvents estimating the abstract density functions $f^{(1)}_{\mathcal{M}}(\cdot|x)$, $x \in W \subseteq \mathbb{R}^n$. Arguably, it is also not an extremely restrictive assumption, at least compared to e.g. independent marking. Moreover, under this assumption, by setting $t(\cdot) = 1$, the summary statistic above reduces to the second-order reduced moment measure $\mathcal{K}(\cdot)$ of the ground process (Møller and Waagepetersen 2004, Definition 4.5), which in turn reduces to the inhomogeneous K-function of Baddeley et al. (2000) when A is given by the closed origin-centred r-ball in $\mathbb{R}^n$.

Ghorbani et al. (2021) further suggest how one may construct inhomogeneous mark-correlation type statistics from $\int_{\mathcal{M}^2} t(m, m')\gamma_{\mathcal{M}}^{(2)}(m, m'|0, x')\nu_{\mathcal{M}}(dm')\nu_{\mathcal{M}}(dm)$, which reduce to the classical ones under stationarity. To shed some light on this connection, in the stationary setting, in addition to the reference measure being given by the common mark distribution, we have that $\lambda_{\mathcal{S}}^{(1)}(\cdot) \equiv \lambda_{\mathcal{S}} > 0$ is constant. Hereby,

$$\begin{aligned}\lambda_{\mathcal{S}}^2\mathcal{K}_t(b(0, r)) &= \int_{b(0,r)} \int_{\mathcal{M}^2} t(m, m') f_{\mathcal{M}}^{(2)}(m, m'|u)\nu_{\mathcal{M}}(dm)\nu_{\mathcal{M}}(dm')\lambda_{\mathcal{S}}^{(2)}(u)du \\ &= \int_{b(0,r)} \mathbb{E}_{0,u}\left[t(m(0), m(u))\right]\lambda_{\mathcal{S}}^{(2)}(u)du,\end{aligned}$$

where, because of the stationarity, $f_{\mathcal{M}}^{(2)}(m, m'|u) = f_{\mathcal{M}}^{(2)}(\cdot|0, u)$ and $\lambda_{\mathcal{S}}^{(2)}(u) = \lambda_{\mathcal{S}}^{(2)}(0, u)$ depend only on the separation vector u rather than the exact locations of the two associated ground points. If X is also isotropic, the inner integral/expectation corresponds to the mark-correlation type statistics reviewed in the paper under discussion. Note that, under these assumptions, Eckardt et al. (2023) look closer at specific choices for the test function in the setting of functional marked point processes. As the review part of the paper under discussion clearly indicates, in practice, the art here lies in making sensible choices for the test function t.

Motivated by the summary statistic of Ghorbani et al. (2021) and D'Angelo et al. (2023) introduced a general family of weighted local summary statistics, which generate e.g. the estimators of the summary statistics of Ghorbani et al. (2021). More specifically, given some test function $\tilde{t}$ on $(\mathbb{R}^n \times \mathcal{M})^2$, in the specific case of second order summary statistics, their local statistics have the form

$$\mathcal{L}_2((x, m), X; \tilde{t}, \tilde{\lambda}) = \sum_{(x',m')\in X} \frac{\tilde{t}((x, m), (x', m'))}{\tilde{\lambda}(x, m)\tilde{\lambda}(x', m')}, \qquad (x, m) \in \mathbb{R}^n \times \mathcal{M}.$$

Here, $\tilde{\lambda}$ is either the true intensity function or an estimator thereof. They verified that by summing up all $\mathcal{L}_2((x, m), X\setminus\{(x, m)\}; \tilde{t}, \tilde{\lambda})$, $(x, m) \in X$, different choices for $\tilde{t}$ and $\tilde{\lambda}$ yield the estimators of the second-order summary statistics in Baddeley et al. (2000), Gabriel and Diggle (2009), Cronie and van Lieshout (2015), Cronie and van Lieshout (2016), Iftimi et al. (2019), Ghorbani et al. (2021).

2.6. LINEAR NETWORK SUMMARY STATISTICS

Next, let X be a marked point process on a linear network $\mathcal{S} = L$, equipped with a regular distance metric d_L (Rakshit et al. 2017). Considering the combination of Cronie and van Lieshout (2016) and Cronie et al. (2020), we believe that the paper under discussion's definitions of inhomogeneous higher-order summary statistics for multi-type point processes on linear networks are sound. However, some appropriate conditions need to be imposed on X. Specifically, recalling the definition of IRMS and having a closer look at Cronie and van Lieshout (2016) and Cronie et al. (2020), it seems that we need to impose that the intensity is bounded away from 0 and that

$$\begin{aligned} g^{(n)}((x_1, m_1), \dots, ((x_n, m_n)) &= g_{\mathcal{S}}^{(n)}(x_1, \dots, x_n)\gamma_{\mathcal{M}}^{(n)}(m_1, \dots, m_n|x_1, \dots, x_n) \\ &= \bar{g}_{\mathcal{S}}^{(n)}(d_L(o, x_1), \dots, d_L(o, x_n))\bar{\gamma}_{\mathcal{M}}^{(n)} \\ &\quad (m_1, \dots, m_n|d_L(o, x_1), \dots, d_L(o, x_n)) \\ &= \bar{g}^{(n)}(d_L(o, x_1), \dots, d_L(o, x_n), m_1, \dots, m_n) \end{aligned}$$

for any $n \geq 1$ and any reference point/origin $o \in \mathcal{S}$ on the network. The relationship between $g_{\mathcal{S}}^{(n)}$ and $\bar{g}_{\mathcal{S}}^{(n)}$ is precisely the intensity reweighted moment pseudostationarity (IRMPS) condition of Cronie et al. (2020). Moreover, $\bar{\gamma}_{\mathcal{M}}^{(n)}$ tells us that the mark density ratio $\gamma_{\mathcal{M}}^{(n)}$, spatially, only depends on the distances between each of $x_1, \dots, x_n \in \mathcal{S}$ and the origin. In other words, this should be the appropriate version of marked IRMPS. Next, for any (Borel) $D \subseteq \mathcal{M}$, consider

$$f_D^u(r, X) = \prod_{(x,m)\in X} \left(1 - \frac{\bar{\lambda}_D \mathbf{1}\{d_L(u, x) \leq r\}\mathbf{1}\{m \in D\} w_{d_L}(u, d_L(u, x))}{\lambda^{(1)}(x, m)}\right),$$

where w_{d_L} is the geometry weight function appearing in Cronie et al. (2020, Sect. 2.3.1). As in the setting of Cronie and van Lieshout (2016), we may also replace $\bar{\lambda}_D$ by any positive but smaller value, e.g. $\bar{\lambda}_{\mathcal{M}}$, in the definition of $f_D^u(r, X)$. Inspection of the definitions and results in Cronie and van Lieshout (2016) and Cronie et al. (2020) reveals that $G_{d_L}^{CD}(r) = 1 - \nu_{\mathcal{M}}(C)^{-1} \int_C \mathbb{E}^!_{(u,m')}[f_D^u(r, X)]\nu_{\mathcal{M}}(dm')$, $F_{d_L}^D(r) = 1 - \mathbb{E}[f_D^u(r, X)]$ and $J_{d_L}^{CD}(r) = (1 - G_{d_L}^{CD}(r))/(1 - F_{d_L}^D(r))$ are linear network versions of the marked inhomogeneous summary statistics of Cronie and van Lieshout (2016). In addition, truncation of the series expansion of $G_{d_L}^{CD}(r)$ gives rise to a linear network analogue of the marked inhomogeneous K-function of Cronie et al. (2020). It may be verified that, due to the proposed form of marked IRMPS, the summary statistics remain the same for (almost) all u that we (Palm) condition on. This, in turn, allows us to define proper estimators, as sums over the points of X (restricted to a spatial study region). In the multi-type setting, if we set $C = \{i\}$ and $D = \{j\}$, and let $\nu_{\mathcal{M}}$ be given by the counting measure on $\mathcal{M} = \{1, \dots, k\}$, then $\lambda^{(1)}(x, i) = \lambda_i(x)$, i.e. the intensity function of X_i, $i = 1, \dots, k$, and $\bar{\lambda}_D = \bar{\lambda}_j = \inf_{x\in\mathbb{R}^n} \lambda_j(x)$. Hence, these summary statistics truly generalise the multivariate summary statistics proposed in the paper under discussion.

Funding Open access funding provided by Chalmers University of Technology.

[Accepted February 2024. Published Online March 2024.]

REFERENCES

Baddeley AJ, Møller J, Waagepetersen R (2000) Non-and semi-parametric estimation of interaction in inhomogeneous point patterns. Stat Neerl 54(3):329–350

Billiot J-M, Coeurjolly J-F, Drouilhet R (2008) Maximum pseudolikelihood estimator for exponential family models of marked Gibbs point processes. Electron J Stat 2:234–264

Coeurjolly J-F, Lavancier F (2013) Residuals and goodness-of-fit tests for stationary marked Gibbs point processes. J R Stat Soc Ser B Stat Methodol 75(2):247–276

Coeurjolly J-F, Dereudre D, Drouilhet R, Lavancier F (2012) Takacs–Fiksel method for stationary marked Gibbs point processes. Scand J Stat 39(3):416–443

Cronie O (2012) Likelihood inference for a functional marked point process with Cox-Ingersoll-Ross process marks. arXiv preprint arXiv:1210.2071

Cronie O, van Lieshout MNM (2015) A J-function for inhomogeneous spatio-temporal point processes. Scand J Stat 42(2):562–579

Cronie O, van Lieshout MNM (2016) Summary statistics for inhomogeneous marked point processes. Ann Inst Stat Math 68(4):905–928

Cronie O, Särkkä A (2011) Some edge correction methods for marked spatio-temporal point process models. Comput Stat Data Anal 55(7):2209–2220

Cronie O, Nyström K, Yu J (2013) Spatiotemporal modeling of Swedish scots pine stands. For Sci 59(5):505–516

Cronie O, Moradi M, Mateu J (2020) Inhomogeneous higher-order summary statistics for point processes on linear networks. Stat Comput 30(5):1221–1239

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

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

D'Angelo N, Adelfio G, Mateu J, Cronie O (2023) Local inhomogeneous weighted summary statistics for marked point processes. J Comput Graph Stat 1–27 (just-accepted)

Dvořák J, Mrkvička T, Mateu J, Gonzalez JA (2022) Nonparametric testing of the dependence structure among points-marks-covariates in spatial point patterns. Int Stat Rev 90(3):592–621

Eckardt M, Comas C, Mateu J (2023) Summary characteristics for multivariate function-valued spatial point process attributes. arXiv preprint arXiv:2307.05101

Gabriel E, Diggle PJ (2009) Second-order analysis of inhomogeneous spatio-temporal point process data. Stat Neerl 63(1):43–51

Ghorbani M, Cronie O, Mateu J, Yu J (2021) Functional marked point processes: a natural structure to unify spatio-temporal frameworks and to analyse dependent functional data. TEST 30:529–568

Goulard M, Särkkä A, Grabarnik P (1996) Parameter estimation for marked Gibbs point processes through the maximum pseudo-likelihood method. Scand J Stat 365–379

Grabarnik P, Särkkä A (2009) Modelling the spatial structure of forest stands by multivariate point processes with hierarchical interactions. Ecol Model 220(9–10):1232–1240

Högmander H, Särkkä A (1999) Multitype spatial point patterns with hierarchical interactions. Biometrics 55(4):1051–1058

Iftimi A, Cronie O, Montes F (2019) Second-order analysis of marked inhomogeneous spatiotemporal point processes: applications to earthquake data. Scand J Stat 46(3):661–685

Lavancier F, Le Guével R (2021) Spatial birth-death-move processes: basic properties and estimation of their intensity functions. J R Stat Soc Ser B Stat Methodol 83(4):798–825

Lotwick H, Silverman B (1982) Methods for analysing spatial processes of several types of points. J R Stat Soc Ser B, Stat Methodol 44(3):406–413

Molchanov IS (2005) Theory of random sets, vol 19. Springer

Møller J, Ghorbani M (2012) Aspects of second-order analysis of structured inhomogeneous spatiotemporal point processes. Stat Neerl 66(4):472–491

Møller J, Waagepetersen R (2004) Statistical inference and simulation for spatial point processes. CRC Press

Mrkvička T, Dvořák J, Gonzalez JA, Mateu J (2021) Revisiting the random shift approach for testing in spatial statistics. Spat Stat 42:100430

Rakshit S, Nair G, Baddeley A (2017) Second-order analysis of point patterns on a network using any distance metric. Spat Stat 22:129–154

Redenbach C, Särkkä A (2013) Parameter estimation for growth interaction processes using spatiotemporal information. Comput Stat Data Anal 57(1):672–683

Reinhart A (2018) A review of self-exciting spatio-temporal point processes and their applications. Stat Sci 33(3):299–318

Renshaw E, Särkkä A (2001) Gibbs point processes for studying the development of spatial-temporal stochastic processes. Comput Stat Data Anal 36(1):85–105

Renshaw E, Comas C, Mateu J (2009) Analysis of forest thinning strategies through the development of space-time growth-interaction simulation models. Stoch Env Res Risk Assess 23:275–288

Särkkä A, Renshaw E (2006) The analysis of marked point patterns evolving through space and time. Comput Stat Data Anal 51(3):1698–1718

van Lieshout MNM (2000) Markov point processes and their applications. Imperial College Press/World Scientific

van Lieshout MNM (2006) A J-function for marked point patterns. Ann Inst Stat Math 58(2):235–259

van Lieshout MNM (2011) A J-function for inhomogeneous point processes. Stat Neerl 65(2):183–201