On the expected value and variance for an estimator of the spatio-temporal product density function

Rodríguez-Cortés, Francisco J. ; Ghorbani, Mohammad; Mateu, Jorge; Stoyan, Dietrich

Publication date:
2014

Document Version
Early version, also known as pre-print

AALBORG UNIVERSITY

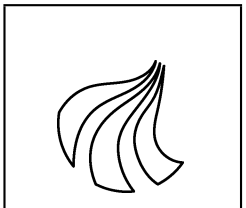
On the expected value and variance for an estimator of the spatio-temporal product density function

by

F. J. Rodríguez-Cortés, M. Ghorbani, J. Mateu and D. Stoyan

R-2014-03 March 2014

Department of Mathematical Sciences
Aalborg University
Fredrik Bajers Vej 7 G ▪ DK - 9220 Aalborg Øst ▪ Denmark
Phone: +45 99 40 99 40 ▪ Telefax: +45 99 40 35 48
URL: http://www.math.aau.dk



ISSN 1399–2503 ▪ On-line version ISSN 1601–7811

On the expected value and variance for an estimator of the spatio-temporal product density function

Francisco J. Rodríguez-Cortés[a], Mohammad Ghorbani[b], Jorge Mateu[a], and **Dietrich Stoyan[c],[1]**

[a] Department of Mathematics, Universitat Jaume I, Castellón, Spain

[b] Department of Mathematics Sciences, Aalborg University, Aalborg, Denmark

[c] Institut für Stochastik, TU Bergakademie Freiberg, Freiberg, Germany

Abstract

Second-order characteristics are used to analyse the spatio-temporal structure of the underlying point process, and thus these methods provide a natural starting point for the analysis of spatio-temporal point process data. We restrict our attention to the spatio-temporal product density function, and develop a non-parametric edge-corrected kernel estimate of the product density under the second-order intensity-reweighted stationary hypothesis. The expectation and variance of the estimator are obtained, and closed form expressions derived under the Poisson case. A detailed simulation study is presented to compare our close expression for the variance with estimated ones for Poisson cases. The simulation experiments show that the theoretical form for the variance gives acceptable values, which can be used in practice. Finally, we apply the resulting estimator to data on the spatio-temporal distribution of invasive meningococcal disease in Germany.

Keywords and Phrases: Edge Correction, Spatio-temporal separability, Second-order product density, Second-order intensity-reweighted stationarity, Variance.

1 Introduction

Spatial and spatio-temporal point patterns are increasingly available in a wide range of scientific settings, such as environmental sciences, climate prediction and meteorology, epidemiology, image analysis, agriculture and astronomy. Today, much attention is paid to spatio-temporal point processes, where each point represents the location and the time of an event, and thus we have data of the form $(\mathbf{u}_i, s_i) \in W \times T \subset \mathbb{R}^2 \times \mathbb{R}$, $i = 1, ..., n$. There are some recent works on spatio-temporal models with focusing on variety of ad-hoc approaches (Diggle (2006); Gabriel and Diggle (2009); Møller and Diaz-Avalos (2010) Gelfand et al. (2010); Diggle (2013)). We consider here processes that are temporally continuous and either spatially continuous or spatially discrete on a sufficiently large support to justify formulating explicitly second-order spatio-temporal tools for the data.

For these processes second-order properties play an important role in the practical analysis of point patterns, in terms of exploratory and modelling strategies. Usually, the K-function and pair correlation function ($g(\cdot)$) are used for model checking (Møller and Ghorbani; 2013) and parameter

[1] All authors contributed equally in this work.

estimation (Møller and Ghorbani; 2012), while the product density is used for explanatory analysis. The form of these functions helps to understand the type of interaction in the point pattern and to find suitable point process models.

In this context, separate analyses of the spatial and the temporal components are of limited value, because the scientific objectives of the analysis are to understand and to model the underlying spatio-temporally interacting stochastic mechanisms. There are basically two ways for modelling spatio-temporal point patterns (Diggle (2006); Daley and Vere-Jones (2008)). The first is descriptive and aims at providing an empirical description of the data, especially from second-order characteristics. The second is mechanistic and aims at constructing parametric point process models by specifying parametric models for the conditional intensity function. Here, we will consider the former and analyses will be based on extensions of the product density to summarize a spatio-temporal point pattern.

The inhomogeneous K-function has been extended to the spatio-temporal setting by Gabriel and Diggle (2009). Second-order characteristics are thus analysed using the spatio-temporal inhomogeneous K-function (STIK-function) or equivalently considering the spatio-temporal pair correlation function under the assumption of second-order intensity re-weighted stationarity (Gabriel and Diggle (2009); Gabriel et al. (2010), Gabriel et al. (2012); Gabriel (2013)). Spatio-temporal separability of the STIK-function has been studied in Møller and Ghorbani (2012). These two functions rely very much upon first-order characteristics which are unknown in practice, and replacing the intensity by an estimate must be made carefully as it may imply bias (Baddeley et al. (2000); Gabriel (2013)). However, the product density does not show this problem, as will be shown in this paper.

Little attention has been paid so far to the first- and second-order moments (expected and variance values) of the second-order properties of spatio-temporal processes. And they are needed for performing statistical inference based on these characteristics. In the spatial context we can only refer to Ripley (1988) who developed variance expressions for a series of estimators of the spatial K-function for the Poisson process.

Then Stoyan et al. (1993) approximated the variance of spatial product densities, and Cressie and Collins (2001a,b) obtained close expressions for the expected and variance values of the local spatial product densities. To the best of our knowledge, nothing has been developed in the spatio-temporal context. In this paper we develop a non-parametric edge-corrected kernel estimate of the product density under the second-order intensity-reweighted stationary hypothesis. We extend the original ideas of Stoyan et al. (1993) to the spatio-temporal case for developing exact and close expressions of the expectation and variance of the proposed estimator. Note that since estimated second-order characteristics deviate from their theoretical counterparts because of statistical fluctuations, error bounds for these functions are important. For example, they are needed to distinguish between statistical fluctuations in an estimated product density function and peaks which are due to real properties of the spatio-temporal point process under study.

Our estimator is accurate in estimating the spatio-temporal product density both under separable and non-separable cases. It is unbiased and we present the close expression of its variance. The simulation experiments show that the formulae derived for this estimator give acceptable values, and thus can be used in practice.

The remainder of the paper is organised as follows. Section 2 provides a theoretical background on the first- and second-order properties of spatio-temporal point processes. In Section 3 we present the product density estimator and its expectation and variance for the general case, and under Poisson processes. Appendix A discusses the corresponding moments of the product density estimator under the hypothesis of separability. We then present some simulation results in Section 4. Section 5 applies our methodology to analysis the spatio-temporal distribution of invasive meningococcal disease in Germany. The paper ends with some final conclusions.

2 Definitions and statistical background

Møller and Ghorbani (2012) discussed the second-order analysis of structured inhomogeneous spatio-temporal point processes. The definitions and notations introduced in that paper are used throughout the present paper. Following them, we consider a spatio-temporal point process with no multiple points as a random countable subset X of $\mathbb{R}^2 \times \mathbb{R}$, where a point $(\mathbf{u}, s) \in X$ corresponds to an event at $\mathbf{u} \in \mathbb{R}^2$ occurring at time $s \in \mathbb{R}$. In practice, we observe n events $\{(\mathbf{u}_i, s_i)\}$ of X within a bounded spatio-temporal region $W \times T \subset \mathbb{R}^2 \times \mathbb{R}$, with area $|W| > 0$, and with length $|T| > 0$. For formal definition of a point process based on measure theory see e.g. (Daley and Vere-Jones; 2008).

For convenience, we introduce the following notations. Let $N(A)$ be the number of events falling in an arbitrary bounded region $A \subset W \times T$; $\Theta_n = \{(\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n) \in X\}$ be a set of n-tuples of events in X; $\int\limits_{B^{\otimes k}} = \int\limits_{B} \ldots \int\limits_{B}$ for k times, where $B = W \times T$.

Assume that X has spatio-temporal nth-order product density function $\rho^{(n)}$, for $n \in \mathbb{N}$. For any non-negative Borel function f defined on $(\mathbb{R}^2 \times \mathbb{R})^{\otimes n}$,

$$\mathbb{E} \sum_{\Theta_n}^{\neq} f((\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)) = \int\limits_{B^{\otimes n}} f((\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)) \times \rho^{(n)}((\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)) \, \mathrm{d}((\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)), \quad (1)$$

where $\sum^{\neq}$ means that we sum over the n pairwise distinct points $(\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)$ (see e.g. (Møller and Waagepetersen; 2004; Illian et al.; 2008; Chiu et al.; 2013)).

2.1 First- and second-order properties

Considering (1), in particular for $n = 1$ and $n = 2$ the n-order product density function is respectively called the intensity function and the second-order product density (hereafter product density) function.

A process for which $\rho(\mathbf{u}, s) = \rho$ for all $(\mathbf{u}, s)$ is called homogeneous or first-order stationary. Further, if $\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) = \rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l)$, the process is called second-order or weak stationary (Ghorbani; 2013).

2.2 Spatial and temporal components

It is assumed that the point process X is orderly, roughly meaning that coincident points cannot occur. That is, any pair of points $(\mathbf{u}, s)$ and $(\mathbf{v}, l)$ of X are distinct, so $\mathbf{u} \neq \mathbf{v}$ and $s \neq l$. We can therefore ignore the case where the spatial and temporal component processes X_{space} and X_{time} have multiple points, and following Møller and Ghorbani (2012) we define them by

$$X_{\text{space}} = \{\mathbf{u} : (\mathbf{u}, s) \in X, s \in T\}, \qquad X_{\text{time}} = \{s : (\mathbf{u}, s) \in X, \mathbf{u} \in W\}.$$

Note that, using this notation, it is clear that X_{space} depends on T, and X_{time} depends on W.

2.2.1 First-order properties

Assume that X has intensity function $\rho(\mathbf{u}, s)$, then

$$\rho_{\text{space}}(\mathbf{u}) = \int\limits_{T} \rho(\mathbf{u}, s) \, \mathrm{d}s \quad \text{and} \quad \rho_{\text{time}}(s) = \int\limits_{W} \rho(\mathbf{u}, s) \, \mathrm{d}\mathbf{u}.$$

Throughout the paper we assume first-order spatio-temporal separability, i.e.

$$\rho(\mathbf{u}, s) = \bar{\rho}_1(\mathbf{u})\bar{\rho}_2(s), \qquad (\mathbf{u}, s) \in \mathbb{R}^2 \times \mathbb{R}, \tag{2}$$

where $\bar{\rho}_1$ and $\bar{\rho}_2$ are non-negative functions.

Considering the hypothesis of first-order spatio-temporal separability,

$$\rho(\mathbf{u}, s) = \frac{\rho_{\text{space}}(\mathbf{u})\rho_{\text{time}}(s)}{\int \rho(\mathbf{u}, s)\,\mathrm{d}(\mathbf{u}, s)}.$$

For a stationary point process X, ρ, ρ_{space} and ρ_{time} are all constant. For non-parametric estimation of $\rho_{\text{space}}(\mathbf{u})$, $\rho_{\text{time}}(s)$ and $\rho(\mathbf{u}, s)$, see Ghorbani (2013).

2.2.2 Second-order properties

Throughout the paper we assume that X is second-order intensity-reweighted stationary (SOIRS), i.e.

$$\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) = \rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l), \qquad (\mathbf{u}, s), (\mathbf{v}, l) \in \mathbb{R}^2 \times \mathbb{R} \tag{3}$$

(Baddeley et al.; 2000; Gabriel and Diggle; 2009; Gabriel; 2013). Further, if the process is isotropic, then $\rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l) = \rho_0^{(2)}(\|\mathbf{u} - \mathbf{v}\|, |s - l|)$ for some non-negative function $\rho_0^{(2)}(\cdot)$, where $\|\cdot\|$ denotes the Euclidean distance in $\mathbb{R}^2$ and $|\cdot|$ denotes the usual distance in $\mathbb{R}$.

Using (1) (with $n = 2$) and (3) we obtain that X_{space} is SOIRS with product density

$$\rho_{\text{space}}^{(2)}(\mathbf{u}, \mathbf{v}) = \rho_{\text{space}}^{(2)}(\mathbf{u} - \mathbf{v}) = \int_T \int_T \rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l)\,\mathrm{d}s\,\mathrm{d}l. \tag{4}$$

Analogously, X_{time} is SOIRS with

$$\rho_{\text{time}}^{(2)}(s, l) = \rho_{\text{time}}^{(2)}(s - l) = \int_W \int_W \rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l)\,\mathrm{d}\mathbf{u}\,\mathrm{d}\mathbf{v}. \tag{5}$$

It will always be clear from the context whether $\rho_{\text{space}}^{(2)}$ is considered to be a function defined on $\mathbb{R}^2 \times \mathbb{R}^2$ or $\mathbb{R}^2$, and whether $\rho_{\text{time}}^{(2)}$ is considered to be a function defined on $\mathbb{R} \times \mathbb{R}$ or $\mathbb{R}$.

2.2.3 Spatio-temporal separability of the product density function

The spatio-temporal product density function is separable if

$$\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) = \bar{\rho}_1^{(2)}(\mathbf{u}, \mathbf{v})\bar{\rho}_2^{(2)}(s, l)$$

for non-negative functions $\bar{\rho}_1^{(2)}$ and $\bar{\rho}_2^{(2)}$. Under the assumption (3) of SOIRS, this hypothesis can be rewritten as

$$\rho^{(2)}(\mathbf{u} - \mathbf{v}, s - l) = \bar{\rho}_1^{(2)}(\mathbf{u} - \mathbf{v})\bar{\rho}_2^{(2)}(s - l), \qquad (\mathbf{u}, s), (\mathbf{v}, l) \in \mathbb{R}^2 \times \mathbb{R}. \tag{6}$$

Considering (4), (5), and (6),

$$\rho_{\text{space}}^{(2)}(\mathbf{u} - \mathbf{v}) = \bar{\rho}_1^{(2)}(\mathbf{u} - \mathbf{v}) \int_T \int_T \bar{\rho}_2^{(2)}(s - l)\,\mathrm{d}s\,\mathrm{d}l, \tag{7}$$

and

$$\rho^{(2)}_{\text{time}}(s-l) = \bar{\rho}^{(2)}_2(s-l) \int\limits_W \int\limits_W \bar{\rho}^{(2)}_1(\mathbf{u}-\mathbf{v})\,\mathrm{d}\mathbf{u}\,\mathrm{d}\mathbf{v}. \tag{8}$$

By substituting (7) and (8) in (6),

$$\rho^{(2)}(\mathbf{u}-\mathbf{v}, s-l) = \frac{\rho^{(2)}_{\text{space}}(\mathbf{u}-\mathbf{v})\rho^{(2)}_{\text{time}}(s-l)}{\int\int \rho^{(2)}(\mathbf{u}-\mathbf{v}, s-l)\,\mathrm{d}(\mathbf{u}, s)\,\mathrm{d}(\mathbf{v}, l)}. \tag{9}$$

As in the spatio-temporal first-order case, equation (9) suggests that

$$\rho^{(2)}((\mathbf{u}, s), (\mathbf{v}, l)) \propto \rho^{(2)}_{\text{space}}(\mathbf{u}, \mathbf{v})\rho^{(2)}_{\text{time}}(s, l).$$

Suppose that $\widehat{\rho^{(2)}}_{\text{space}}(\mathbf{u}-\mathbf{v})$ respective $\widehat{\rho^{(2)}}_{\text{time}}(s-l)$ are estimators of ${\rho^{(2)}}_{\text{space}}(\mathbf{u}-\mathbf{v})$ respective ${\rho^{(2)}}_{\text{time}}(s-l)$. If these are unbiased estimates of the expected number of distinct pairs of events, i.e. $\int_W \widehat{\rho^{(2)}}_{\text{space}}(\mathbf{u}-\mathbf{v})\,\mathrm{d}\mathbf{u}\,\mathrm{d}\mathbf{v} = \int_T \widehat{\rho^{(2)}}_{\text{time}}(s-l)\,\mathrm{d}s\,\mathrm{d}l = n(n-1)$, then the estimate of the spatio-temporal product density function given by

$$\widehat{\rho^{(2)}}(\mathbf{u}-\mathbf{v}, s-l) = \frac{\widehat{\rho^{(2)}}_{\text{space}}(\mathbf{u}-\mathbf{v})\widehat{\rho^{(2)}}_{\text{time}}(s-l)}{n(n-1)},$$

is also a ratio unbiased estimate of the expected number of observed points. See more details in Section 3.

2.3 Relationship between the product density and the K-function

For a SOIRS, isotropic, spatio-temporal point process X, Gabriel and Diggle (2009) extended the inhomogeneous K-function from the spatial to the spatio-temporal case. They defined the spatio-temporal inhomogeneous K-function as

$$K(r,t) = \int \mathbf{1}\left[\|\mathbf{u}\| \le r, |s| \le t\right] g_0(\mathbf{u}, s)\mathrm{d}(\mathbf{u}, s), \qquad r > 0, \quad t > 0, \tag{10}$$

where $\mathbf{1}[\cdot]$ denotes the indicator function, and $g_0(\mathbf{u}, s)$ (with the abuse of the notations $\mathbf{u}$ and s for $\mathbf{u} = \|\mathbf{u}-\mathbf{v}\|$ and $s = |s-l|$) is the spatio-temporal pair correlation function. For a Poisson process, $g_0 = 1$ and $K(r,t) = 2\pi r^2 t$. For an unbiased estimator of the K-function, see Gabriel (2013).

Considering the hypothesis of the first- and second-order spatio-temporal separabilities, for isotropic point process X and for non-negative Borel functions $\bar{K}_1$ and $\bar{K}_2$,

$$K(r,t) = \bar{K}_1(r)\bar{K}_2(t), \qquad r > 0, \quad t > 0. \tag{11}$$

Assume that X is isotropic, and X_{space} and X_{time} have pair correlation functions g_{space} and g_{time} respectively. The corresponding spatial and temporal K-functions are

$$K_{\text{space}}(r) = \int\limits_{||\mathbf{u}|| \le r} g_{\text{space}}(\mathbf{u})\,\mathrm{d}\mathbf{u}, \qquad r > 0, \quad \text{and} \quad K_{\text{time}}(t) = \int\limits_{-t}^{t} g_{\text{time}}(s)\,\mathrm{d}s, \qquad t > 0.$$

Both in the stationary and isotropic case, and in the SOIRS and isotropic case, the spatio-temporal pair correlation function is proportional to the derivative of $K(r,t)$ with respect to r and t. So, in the planar case using (10),

$$g_0(r,t) = \frac{1}{4\pi r}\frac{\partial^2 K(r,t)}{\partial r \partial t}.$$

Thus, for the SOIRS and isotropic point process X,

$$\rho^{(2)}((\mathbf{u},s),(\mathbf{v},l)) = \frac{\rho(\mathbf{u},s)\rho(\mathbf{v},l)}{4\pi r}\frac{\partial^2 K(r,t)}{\partial r \partial t}, \qquad r>0, \quad t>0.$$

Further, under spatio-temporal separability (2) and (11), we have that

$$\rho^{(2)}(r,t) = \frac{1}{c_{\mathrm{SO}}}\left(\frac{\rho_{\mathrm{space}}(\mathbf{u})\rho_{\mathrm{space}}(\mathbf{v})}{2\pi r}\frac{\partial K_{\mathrm{space}}(r)}{\partial r}\right)\left(\frac{\rho_{\mathrm{time}}(s)\rho_{\mathrm{time}}(l)}{2}\frac{\partial K_{\mathrm{time}}(t)}{\partial t}\right). \tag{12}$$

Here

$$c_{\mathrm{SO}} = c^s_{\mathrm{SO}} \times c^t_{\mathrm{SO}},$$

with

$$c^s_{\mathrm{SO}} = \left(\int_W \bar{\rho}^{(2)}(\mathbf{u})\,\mathrm{d}\mathbf{u}\right) \qquad \text{and} \qquad c^t_{\mathrm{SO}} = \left(\int_T \bar{\rho}^{(2)}(s)\,\mathrm{d}s\right),$$

and then

$$c_{\mathrm{SO}} = \left(\iint \rho^{(2)}(\mathbf{u},s)\,\mathrm{d}(\mathbf{u},s)\right),$$

which can be approximated by $n(n-1)$. Hence

$$\rho^{(2)}_{\mathrm{space}}(r) \propto \frac{\rho_{\mathrm{space}}(\mathbf{u})\rho_{\mathrm{space}}(\mathbf{v})}{2\pi r}\frac{\partial K_{\mathrm{space}}(r)}{\partial r}, \quad \rho^{(2)}_{\mathrm{time}}(t) \propto \frac{\rho_{\mathrm{time}}(s)\rho_{\mathrm{time}}(l)}{2}\frac{\partial K_{\mathrm{time}}(t)}{\partial t}.$$

For a stationary and isotropic point process X,

$$\rho^{(2)}(r,t) = \frac{\rho^2}{4\pi r}\frac{\partial^2 K(r,t)}{\partial r \partial t}. \tag{13}$$

Moreover,

$$\rho^{(2)}_{\mathrm{space}}(r) \propto \frac{\rho^2_{\mathrm{space}}}{2\pi r}\frac{\partial K_{\mathrm{space}}(r)}{\partial r}, \qquad \rho^{(2)}_{\mathrm{time}}(t) \propto \frac{\rho^2_{\mathrm{time}}}{2}\frac{\partial K_{\mathrm{time}}(t)}{\partial t}.$$

3 Estimation of the product density function

We avoid estimating the product density by applying numerical differentiation to an estimate of $\rho^2 K(r,t)$. Alternatively, considering that $\rho^2 K(r,t)$ stands for the expected number of ordered pairs of distinct points per unit area of the observation window with pairwise distance and time lag less than r and t, by extending the idea in Stoyan (1987) and Stoyan et al. (1995), we directly estimate the product density using a non-parametric edge-corrected kernel estimate.

A spatio-temporal kernel density estimate of $\rho^2 \partial K(r,t)/\partial r \partial t$ takes the basic form of a smoothed three-dimensional histogram,

$$(|W||T|)^{-1} \sum_{(\mathbf{u},s),(\mathbf{v},l)\in X}^{\neq} \kappa_{\epsilon\delta}(\|\mathbf{u}-\mathbf{v}\|-r, |s-l|-t).$$

We assume that the kernel function $\kappa_{\epsilon\delta}(\cdot,\cdot)$ has the multiplicative form

$$\kappa_{\epsilon\delta}(\|\mathbf{u}-\mathbf{v}\|-r, |s-l|-t) = \kappa_{1\epsilon}\left(\|\mathbf{u}-\mathbf{v}\|-r\right)\kappa_{2\delta}\left(|s-l|-t\right),$$

where $\kappa_{2\delta}$ and $\kappa_{1\epsilon}$ are one-dimensional kernel functions with bandwidths ϵ and δ, respectively.

By extending the idea in Ohser (1983), an edge-corrected kernel estimate of the product density function (13) is given by

$$\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t) = \sum_{(\mathbf{u},s),(\mathbf{v},l)\in X}^{\neq} \frac{\kappa_{1\epsilon}\left(\|\mathbf{u}-\mathbf{v}\|-r\right)\kappa_{2\delta}\left(|s-l|-t\right)}{4\pi r \gamma_W(r)\gamma_T(t)}, \qquad r>\epsilon>0,\ t>\delta>0. \tag{14}$$

Here $\gamma_W(r)$ and $\gamma_T(t)$ are the spatial and temporal set covariance functions, respectively. For a convex region W, a general approximation formula for $\gamma_W(r)$ for small r is given by

$$\gamma_W(r) \approx |W| - \frac{U(W)}{\pi} r,$$

where $U(W)$ is the perimeter of W, and for a small t, $\gamma_T(t) = |T| - t$.

Under the hypothesis of spatio-temporal separability, and considering (9),

$$\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t) \simeq \frac{\widehat{\rho^{(2)}}_{\text{space},\epsilon}(r)\widehat{\rho^{(2)}}_{\text{time},\delta}(t)}{n(n-1)}, \tag{15}$$

with

$$\widehat{\rho^{(2)}}_{\text{space},\epsilon}(r) = \sum_{\mathbf{u},\mathbf{v}\in X_{\text{space}}}^{\neq} \frac{\kappa_{1\epsilon}(\|\mathbf{u}-\mathbf{v}\|-r)}{2\pi r\gamma_W(r)}, \qquad r>\epsilon>0,$$

and

$$\widehat{\rho^{(2)}}_{\text{time},\delta}(t) = \sum_{s,l\in X_{\text{time}}}^{\neq} \frac{\kappa_{2\delta}(|s-l|-t)}{2\gamma_T(t)}, \quad t>\delta>0.$$

3.1 Expectation and variance of the product density estimator

In this section the expectation and variance of the product density estimator (14) is obtained by considering the general case. The corresponding moments of the product density estimator under the hypothesis of separability are developed in Appendix A.

3.1.1 Expectation

Using (1) with $n=2$, the estimator (14) satisfies

$$\begin{aligned}\mathbb{E}\left[\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right] &= \int\int \frac{\kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{y}\|-r)\kappa_{2\delta}(|\xi-\eta|-t)}{4\pi\gamma_W(r)\gamma_T(t)r}\rho^{(2)}(\|\mathbf{x}-\mathbf{y}\|,|\xi-\eta|)\mathrm{d}(\mathbf{x},\xi)\mathrm{d}(\mathbf{y},\eta)\\ &= \int_{-r/\epsilon}^{\infty}\int_{-t/\delta}^{\infty} \frac{\kappa_1(u)\kappa_2(v)\gamma_W(r+\epsilon u)\gamma_T(t+\delta v)}{r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(r+\epsilon u, t+\delta v)(r+\epsilon u)\mathrm{d}u\,\mathrm{d}v\end{aligned} \tag{16}$$

The detailed proof is as follows. By applying formula (1) and the Fubini's theorem to (14), we have that

$$
\begin{aligned}
\mathbb{E}\left[\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right] &= \mathbb{E}\sum_{(\mathbf{u},s),(\mathbf{v},l)\in X}^{\neq} \frac{\mathbf{1}_W(\mathbf{u})\mathbf{1}_W(\mathbf{v})\mathbf{1}_T(s)\mathbf{1}_T(l)\kappa_{1\epsilon}\left(\|\mathbf{u}-\mathbf{v}\|-r\right)\kappa_{2\delta}\left(|s-l|-t\right)}{4\pi r\gamma_W(r)\gamma_T(t)} \\
&= \int\limits_{W\times T}\int\limits_{W\times T} \frac{\kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{y}\|-r)\kappa_{2\delta}(|\xi-\eta|-t)}{4\pi r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(\|\mathbf{x}-\mathbf{y}\|,|\xi-\eta|)\,\mathrm{d}(\mathbf{x},\xi)\,\mathrm{d}(\mathbf{y},\eta) \\
&= \int\limits_{(W-h_1)\times(T-h_2)}\int\limits_{W\times T} \frac{\kappa_{1\epsilon}(\|\mathbf{h_1}\|-r)\kappa_{2\delta}(|h_2|-t)}{4\pi r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(\|\mathbf{h_1}\|,|h_2|)\,\mathrm{d}(\mathbf{h_1},h_2)\,\mathrm{d}(\mathbf{y},\eta) \\
&= \int\limits_{\mathbb{R}^2\times\mathbb{R}} \frac{\kappa_{1\epsilon}(\|\mathbf{h_1}\|-r)\kappa_{2\delta}(|h_2|-t)\gamma_W(h_1)\gamma_T(h_2)}{4\pi r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(\|\mathbf{h_1}\|,|h_2|)\,\mathrm{d}(\mathbf{h_1},h_2) \\
&= \int\limits_0^\infty\int\limits_{\mathbb{R}} \frac{\kappa_{1\epsilon}(R-r)\kappa_{2\delta}(|h_2|-t)\gamma_W(R)\gamma_T(h_2)}{2r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(R,|h_2|)\,R\,\mathrm{d}R\,\mathrm{d}h_2 \\
&= \int\limits_{-r/\epsilon}^\infty\int\limits_0^\infty \frac{\kappa_1(u)\kappa_2((h_2-t)/\delta)\gamma_W(r+u\epsilon)\gamma_T(h_2)}{2\delta r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(r+u\epsilon,h_2)(r+u\epsilon)\,\mathrm{d}u\,\mathrm{d}h_2 \\
&+ \int\limits_{-r/\epsilon}^\infty\int\limits_{-\infty}^0 \frac{\kappa_1(u)\kappa_2((-h_2-t)/\delta)\gamma_W(r+u\epsilon)\gamma_T(h_2)}{2\delta r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(r+u\epsilon,-h_2)(r+u\epsilon)\mathrm{d}u\,\mathrm{d}h_2 \\
&= \int\limits_{-r/\epsilon}^\infty\int\limits_{-t/\delta}^\infty \frac{\kappa_1(u)\kappa_2(v_1)\gamma_W(r+u\epsilon)\gamma_T(\delta v_1+t)}{2r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(r+u\epsilon,\delta v_1+t)(r+u\epsilon)\mathrm{d}u\,\mathrm{d}v_1 \\
&+ \int\limits_{-r/\epsilon}^\infty\int\limits_{-t/\delta}^\infty \frac{\kappa_1(u)\kappa_2(v_2)\gamma_W(r+u\epsilon)\gamma_T(\delta v_2+t)}{2r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(r+u\epsilon,\delta v_2+t)(r+u\epsilon)\mathrm{d}u\,\mathrm{d}v_2 \\
&= \int\limits_{-r/\epsilon}^\infty\int\limits_{-t/\delta}^\infty \frac{\kappa_1(u)\kappa_2(v)\gamma_W(r+u\epsilon)\gamma_T(\delta v+t)}{r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(r+u\epsilon,\delta v+t)(r+u\epsilon)\mathrm{d}u\,\mathrm{d}v
\end{aligned}
$$

If (r,t) is a continuity point of $\rho^{(2)}(r,t)$, then

$$\lim_{(\epsilon,\delta)\to(0,0)}\mathbb{E}\left[\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right]=\rho^{(2)}(r,t).$$

Hence, $\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)$ is an approximately unbiased estimator for the spatio-temporal product density.

3.1.2 Variance

The variance of the product density estimator (14) can be obtained by the direct application of the extended Campbell's Theorem (Illian et al.; 2008; Chiu et al.; 2013) for the spatio-temporal case.

In particular, we have

$$\mathbb{E}\left[\left(\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right)^2\right] = \frac{(c(r,t))^2}{16}\left[4E_1(B) + 2E_2(B) + E_3(B)\right] \tag{17}$$

with

$$c(r,t) = \frac{1}{\pi r \gamma_W(r)\gamma_T(t)}$$

and

$$\begin{aligned} E_1(B) = \int_{B^{\otimes 3}} & \kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{y}\|-r)\kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{z}\|-r)\kappa_{2\delta}(|\xi-\eta|-t)\kappa_{2\delta}(|\xi-\zeta|-t) \\ & \times \rho^{(3)}((\|\mathbf{x}-\mathbf{y}\|,|\xi-\eta|),(\|\mathbf{x}-\mathbf{z}\|,|\xi-\zeta|))\mathrm{d}(\mathbf{x},\xi)\mathrm{d}(\mathbf{y},\eta)\mathrm{d}(\mathbf{z},\zeta), \end{aligned}$$

$$E_2(B) = \int_{B^{\otimes 2}} \kappa_{1\epsilon}^2(\|\mathbf{x}-\mathbf{y}\|-r)\kappa_{2\delta}^2(|\xi-\eta|-t)\rho^{(2)}(\|\mathbf{x}-\mathbf{y}\|,|\xi-\eta|)\mathrm{d}(\mathbf{x},\xi)\mathrm{d}(\mathbf{y},\eta),$$

$$\begin{aligned} E_3(B) = \int_{B^{\otimes 4}} & \kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{y}\|-r)\kappa_{1\epsilon}(\|\mathbf{z}-\mathbf{w}\|-r)\kappa_{2\delta}(|\xi-\eta|-t)\kappa_{2\delta}(|\zeta-\gamma|-t) \\ & \times \rho^{(4)}((\|\mathbf{x}-\mathbf{y}\|,|\xi-\eta|),(\|\mathbf{x}-\mathbf{z}\|,|\xi-\zeta|),(\|\mathbf{x}-\mathbf{w}\|,|\xi-\eta|)) \\ & \times \mathrm{d}(\mathbf{x},\xi)\mathrm{d}(\mathbf{y},\eta)\mathrm{d}(\mathbf{z},\zeta)\mathrm{d}(\mathbf{w},\gamma). \end{aligned}$$

Finding an expansion for the variance in terms of (ϵ,δ) will require knowledge of the form of the third and fourth-order product density function for a given point process model.

3.2 Expectation and variance under Poisson processes

3.2.1 Expectation

For a Poisson process with intensity ρ, the nth-order product density $\rho^{(n)}$ is equal to ρ^n, so using (16) when $(\epsilon,\delta)\to(0,0)$,

$$\begin{aligned} \mathbb{E}\left[\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right] = \int_{-r/\epsilon}^{\infty}\int_{-t/\delta}^{\infty} & \frac{\kappa_1(u)\kappa_2(v)\gamma_W(r+u\epsilon)\gamma_T(t+\delta v)}{r\gamma_W(r)\gamma_T(t)} \\ & \times \rho^{(2)}(r+u\epsilon,t+\delta v)(r+u\epsilon)\mathrm{d}u\,\mathrm{d}v = \rho^2, \end{aligned} \tag{18}$$

if the lower bound for the value of κ_1 and κ_2 are larger than $-r/\epsilon$ and $-t/\delta$, respectively.

3.2.2 Variance

Considering (17), and the unbiasedness property of the product density estimator,

$$\mathbb{V}\mathrm{ar}\left[\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right] = \frac{(c(r,t))^2}{16}\left[4\rho^3 S_1 + 2\rho^2 S_2\right], \tag{19}$$

where

$$S_1 = \int\limits_{B^{\otimes 3}} \kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{y}\|-r)\kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{z}\|-r)\kappa_{2\delta}(|\xi-\eta|-t)\kappa_{2\delta}(|\xi-\zeta|-t)\,\mathrm{d}(\mathbf{x},\xi)\mathrm{d}(\mathbf{y},\eta)\mathrm{d}(\mathbf{z},\zeta)$$
$$= \int\limits_{B} \left\{ \int\limits_{B} \kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{y}\|-r)\kappa_{2\delta}(|\xi-\eta|-t)\,\mathrm{d}(\mathbf{y},\eta) \right\}^2 \mathrm{d}(\mathbf{x},\xi) = S_1^s S_1^t$$

and

$$S_2 = \int\limits_{B^{\otimes 2}} \kappa_{1\epsilon}^2(\|\mathbf{x}-\mathbf{y}\|-r)\kappa_{2\delta}^2(|\xi-\eta|-t)\mathrm{d}(\mathbf{x},\xi)\mathrm{d}(\mathbf{y},\eta) = S_2^s S_2^t.$$

Here,

$$S_1^s = \int\limits_{W} \left\{ \int\limits_{W} \kappa_{1\epsilon}(\|\mathbf{x}-\mathbf{y}\|-r)\,\mathrm{d}\mathbf{y} \right\}^2 \mathrm{d}\mathbf{x}, \qquad S_1^t = \int\limits_{T} \left\{ \int\limits_{T} \kappa_{2\delta}(|\xi-\eta|-t)\,\mathrm{d}\eta \right\}^2 \mathrm{d}\xi,$$

$$S_2^s = \int\limits_{W^{\otimes 2}} \kappa_{1\epsilon}^2(\|\mathbf{x}-\mathbf{y}\|-r)\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} \quad \text{and} \quad S_2^t = \int\limits_{T^{\otimes 2}} \kappa_{2\delta}^2(|\xi-\eta|-t)\mathrm{d}\xi\mathrm{d}\eta.$$

For the spatial case, and using the Epanechnikov kernel, Stoyan et al. (1993) showed that

$$S_2^s = \frac{6}{5\epsilon}\left(|W|\pi r - U(W)\left(\frac{\epsilon^2}{7}+r^2\right)\right) \quad \text{and} \quad S_1^s = 4\pi^2 r^2\big(|W|-A\big) + 4(r+\epsilon)^2(\pi-1)^2 A,$$

where $A = U(W)(r+\epsilon) - 4(r+\epsilon)^2$.

For the temporal case, by using the uniform kernel, it is easy to show that

$$S_2^t = \frac{|T|}{\delta} \quad \text{and} \quad S_1^t = 4|T| - 8(t+\delta) + \frac{128}{3}t^2(t+\delta).$$

By combining the above expressions, an estimate of the variance of the product density estimator is obtained. In practice we substitute ρ by its estimate $\hat{\rho} = \frac{N(W\times T)}{|W||T|}$.

4 Simulation study

The spatio-temporal product density function is of interest as it can be used to discriminate among several spatio-temporal point process models. For example, for a Poisson process $\mathbb{E}[\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)] = \rho^2$ as we have shown previously in (18). Values of the spatio-temporal surface of the product density function larger than the values of the plane ρ^2, indicate that the interevent distances around (r,t) are relatively more frequent compared to those in a Poisson process, which is typical of a cluster process, and conversely, values of the spatio-temporal surface of $\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)$ smaller than the values of the plane ρ^2 indicate that the corresponding distances are rare and this is typical of an inhibition process. The product density function can take all values between zero and infinity.

We conducted a simulation experiment to analyse the behaviour of our estimator of the second-order spatio-temporal product density function under random Poisson structures. In addition, as we have developed close expressions for the variance under the Poisson case, we use them to generate the corresponding confidence surfaces. We considered the volume $W \times T = [0, 10]^2 \times [0, 10]$ and simulated spatio-temporal point patterns with a varying expected number of points $\mathbb{E}[N(W \times T)] = n = 100, 200, 300$. We considered $N_{\text{sim}} = 100$ repetitions per pattern and scenario. The work has been implemented in R, and has used the stpp package Gabriel et al. (2012). We used a fine grid for each spatial and temporal distances u and v spanning the sequence starting from $u > \epsilon > 0$ to 2.50 and $v > \delta > 0$ to 2.50 with small increments of distances. In the spatial case, Fiksel (1988) suggested the use of the Epanechnikov kernel with bandwidth parameter $\epsilon = 0.1\sqrt{5/\rho}$. In practice, we use the dpik function in kernsmooth package to obtain the bandwidth (Wand and Ripley; 2013) based on the distances between the spatial locations of the process. For the temporal case the uniform kernel is used, where again we calculate the bandwidth δ using the dpik function based on the time lag between the temporal instants of the process. Note that the product density function was evaluated for any scenario and repetition over the same spatio-temporal grid.

Table 1 shows some descriptive measures of the second-order spatio-temporal product density kernel estimator for homogeneous Poisson process under different expected number of points. The homogeneous Poisson processes are simulated using the rpp function in stpp package, with constant intensity. The spatial and temporal bandwidths are estimated for each one hundred repetitions. Table 1 displays the average optimal bandwidths for each sample size (named "Est" in Table 1). We also show the behaviour of the product density under two other fixed bandwidth values designed to overestimate and underestimate the optimal values (named "Fix" in Table 1). From all possible grid cells, in Table 1 we have only shown the descriptive measures for some particular values of (r, t), for comparison purposes.

Table 1 also shows the theoretical product density under a Poisson case ($\rho^{(2)} = \widehat{\rho^2}$), together with the estimated average surface ($\widehat{(\rho^{(2)})}$), ($Q_{5\%}(\widehat{\rho^{(2)}})$) and ($Q_{95\%}(\widehat{\rho^{(2)}})$) are the 5% and 95% sample quantile values. Note that, we have estimated $\rho^{(2)}$ by $\frac{n(n-1)}{(|W||T|)^2}$ which is an unbiased estimator in the case of Poisson process (Stoyan and Stoyan; 1994). In terms of variances, we present the average approximate theoretical standard deviation surface ($\sigma(\widehat{\rho^{(2)}})$) together with the average empirical standard deviation surface values ($\widehat{\sigma}(\widehat{\rho^{(2)}})$).

Table 1 shows the results for only three selected cells over the fine grid of spatial and temporal distances to save space. We note that in general the difference between the estimated product density and the theoretical one is smaller when using the estimated bandwidth using dpik, compared with those cases where we use some other fixed values for the bandwidth. In addition, the variances coming from our theoretical developed expression are in the same order of the empirical variance for the selected cells, and even lower for many other cells.

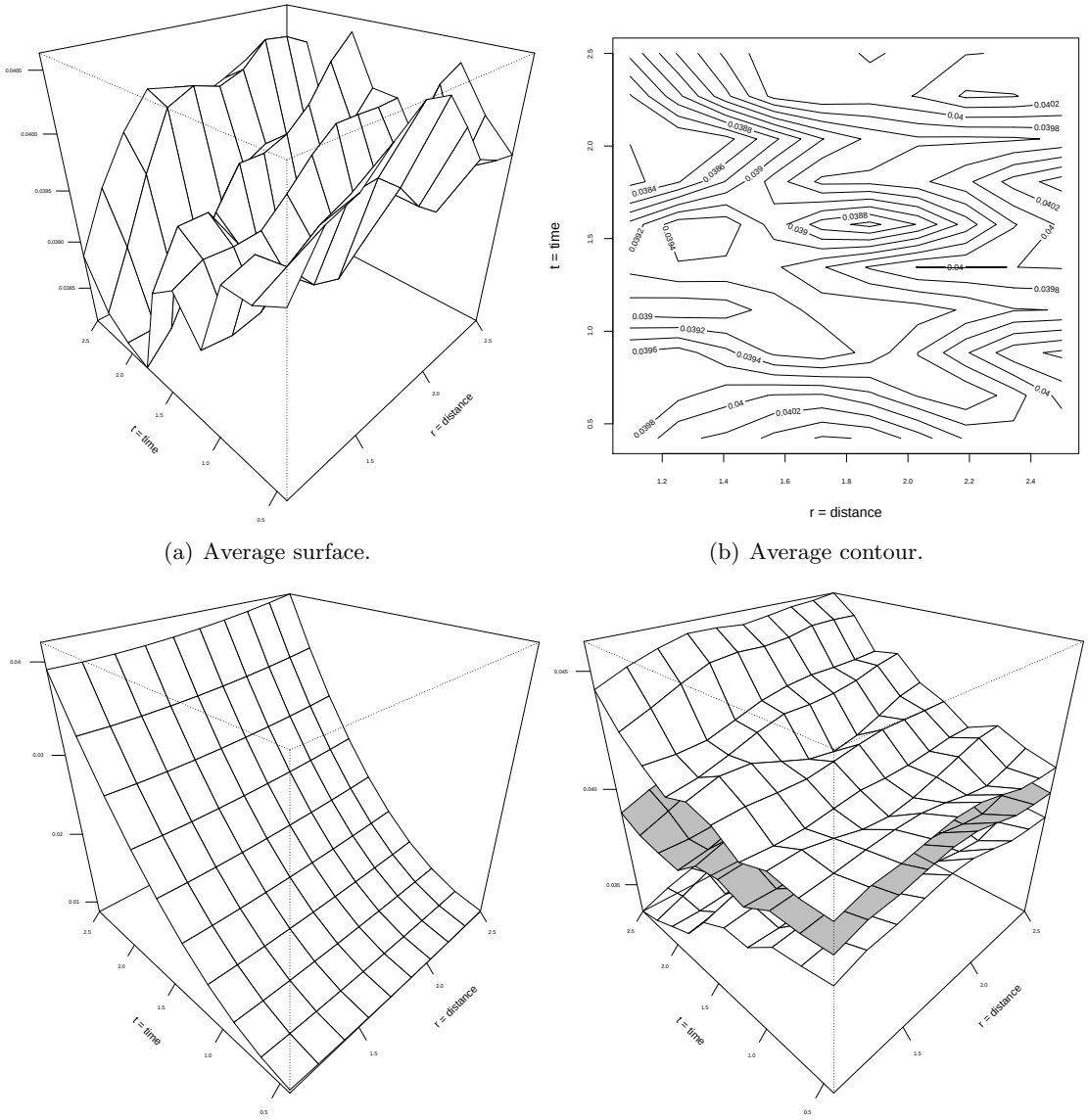
The estimated product density function over the whole grid is depicted in Figures 1, for the case $n = 200$, $\epsilon = 0.7383$ and $\delta = 0.2466$. In this Figure: The top panels show the Monte Carlo mean estimate of $\widehat{\rho^{(2)}}_{\epsilon,\delta}(r, t)$. Clearly $\widehat{\rho^{(2)}}_{\epsilon,\delta}(r, t)$ is an approximate unbiased estimator of $\rho^{(2)}(r, t)$. Bottom left panel shows the Monte Carlo mean estimate of standard deviation. Bottom right panel shows the confidence surfaces under a Poisson process based on the estimated product density (constant surface with value 0.0398) and two standard deviations calculated using the closed form expression of the variance in Section 3.2.2. We have obtained the same results for the cases $n = 100$ and $n = 300$, but the plots are omitted here.

n	**Type**	ϵ	δ	r	t	$\rho^{(2)} = \widehat{\rho^2}$	$Q_{5\%}(\widehat{\rho^{(2)}})$	$\widehat{\rho^{(2)}}$	$Q_{95\%}(\widehat{\rho^{(2)}})$	$\sigma(\widehat{\rho^{(2)}})$	$\widehat{\sigma}(\widehat{\rho^{(2)}})$
100	**Fix.**	0.70	0.15	1.1610	0.6192	0.0099	0.0097	0.0103	0.0160	0.0039	0.0030
				1.6631	1.3245	0.0099	0.0097	0.0104	0.0162	0.0063	0.0034
				2.1653	2.0298	0.0099	0.0097	0.0100	0.0157	0.0115	0.0032
	Est.	0.9936	0.3841	1.1610	0.6192	0.0099	0.0101	0.0102	0.0145	0.0038	0.0025
				1.6631	1.3245	0.0099	0.0099	0.0101	0.0149	0.0069	0.0027
				2.1653	2.0298	0.0099	0.0099	0.0099	0.0154	0.0128	0.0028
	Fix.	1.20	0.50	1.1610	0.6192	0.0099	0.0101	0.0100	0.0140	0.004	0.0023
				1.6631	1.3245	0.0099	0.0100	0.0100	0.0144	0.0075	0.0026
				2.1653	2.0298	0.0099	0.0098	0.0099	0.0151	0.0138	0.0027
200	**Fix.**	0.48	0.07	1.1610	0.6192	0.0398	0.0405	0.0407	0.0576	0.0113	0.0107
				1.6631	1.3245	0.0398	0.0414	0.0413	0.0584	0.0165	0.0105
				2.1653	2.0298	0.0398	0.0398	0.0408	0.0582	0.0293	0.0100
	Est.	0.7383	0.2466	1.1610	0.6192	0.0398	0.0399	0.0403	0.0525	0.0093	0.0067
				1.6631	1.3245	0.0398	0.0405	0.0402	0.0518	0.0167	0.007
				2.1653	2.0298	0.0398	0.0396	0.0402	0.0510	0.0313	0.0069
	Fix.	0.90	0.36	1.1610	0.6192	0.0398	0.0395	0.0398	0.0509	0.0096	0.0062
				1.6631	1.3245	0.0398	0.0392	0.0399	0.0508	0.0178	0.0066
				2.1653	2.0298	0.0398	0.0396	0.0400	0.0497	0.0333	0.0064
300	**Fix.**	0.42	0.03	1.1610	0.6192	0.0897	0.0923	0.0935	0.1305	0.0248	0.0238
				1.6631	1.3245	0.0897	0.0914	0.0944	0.1394	0.0326	0.0236
				2.1653	2.0298	0.0897	0.0946	0.0948	0.1293	0.0546	0.0196
	Est.	0.6093	0.1862	1.1610	0.6192	0.0897	0.0934	0.0933	0.1144	0.0166	0.0135
				1.6631	1.3245	0.0897	0.0941	0.0947	0.1187	0.0295	0.0157
				2.1653	2.0298	0.0897	0.0942	0.0953	0.1219	0.0556	0.0144
	Fix.	0.81	0.34	1.1610	0.6192	0.0897	0.0923	0.0926	0.1124	0.0171	0.0122
				1.6631	1.3245	0.0897	0.0931	0.0936	0.1160	0.0319	0.0137
				2.1653	2.0298	0.0897	0.0929	0.0944	0.1161	0.0600	0.0135

Table 1. Descriptive measures of the estimation of the second-order spatio-temporal product density under the Poisson case.

5 Invasive Meningococcal Disease (IMD): Second-order analysis

Meyer et al. (2012) quantified the transmission dynamics of the two most common meningococcal antigenic sequence types observed in Germany between 2002 and 2008. The conditional intensity function was modelled as a superposition of additive and multiplicative components in space and time. The Invasive Meningococcal Disease (IMD) is a known human disease which involves meningitis (50% of cases), septicemia (5% to 20%) and/or pneumonia (5% to 15%) caused by the infection with the bacterium Neisseria meningitidis. Meningococci can be transmitted airborne or by other mucous secretions from infected humans. The risk of contracting IMD is much higher inside the



(a) Average surface.

(b) Average contour.

(c) Average standard deviation surface.

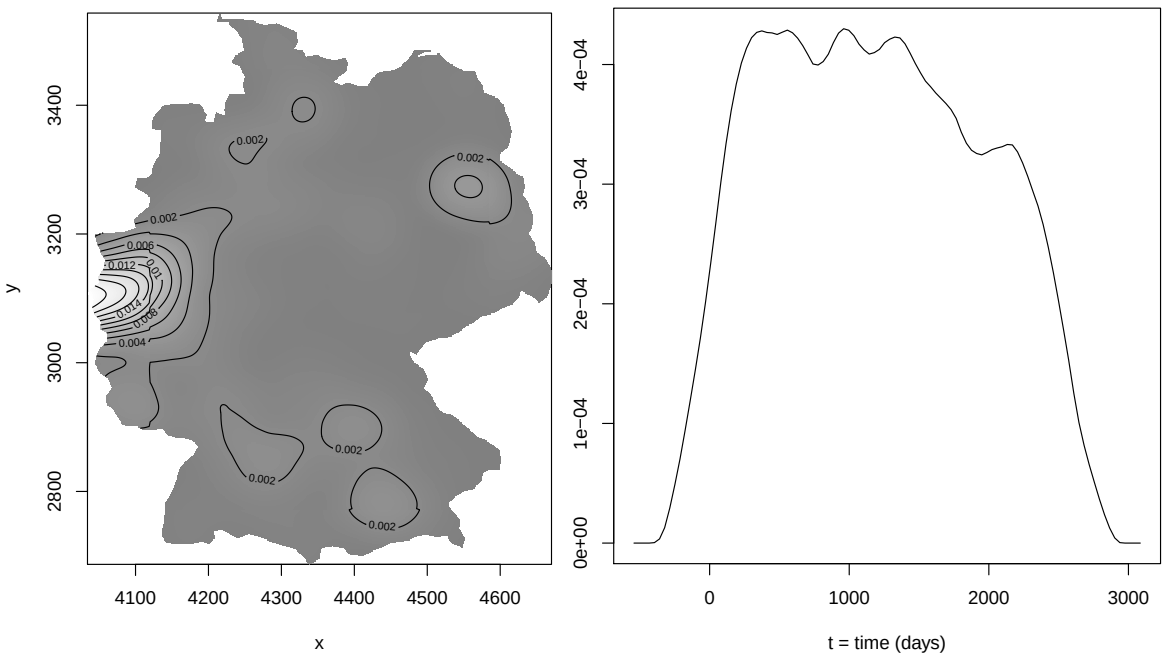
(d) Average confidence surfaces.

Figure 1: Statistical properties of the second-order spatio-temporal product density kernel estimator under Poisson point process with expected number of points $n = 200$, $\widehat{\rho^2} = 0.0398$, $\epsilon = 0.7383$ and $\delta = 0.2466$.

household of an infected person, and the risk of secondary infections is highest during the first few days. Meyer et al. (2012) claim that most meningococci are commensal in humans, but only a few isolates are virulent and cause invasive disease.

The area of Germany is 357603 km^2 with a perimeter of 6146 km. The IMD dataset consists of the spatio-temporal reports of 636 cases of IMD caused by two specific meningococcal finetypes in which the times are given by 2569 days over the 7-year period, so the temporal region is defined as $T = [0, 2569]$. Figure 2 shows the estimated spatial intensity (a) and estimated temporal intensity

(b). In the purely spatial case, this figure shows clearly the inhomogeneity condition of IMD, with a notorious high intensity of points per km^2 in the western border of Germany, and some lower intensity (but noticeable concentrations) near the north-eastern and southern borders.



(a) Estimated spatial intensity.

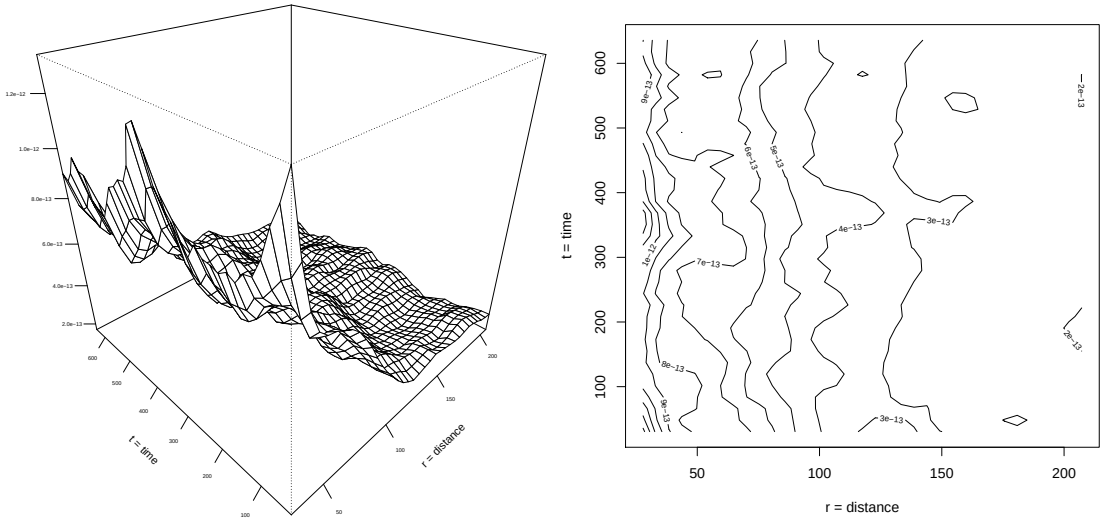
(b) Estimated temporal intensity.

Figure 2: Estimated spatial and temporal intensities for the IMD dataset.

Figure 3 shows the surface of the estimated product density using $\epsilon = 13.9686$ km and $\delta = 28$ days. This figure shows large values for small spatial and temporal distances, which is a typical behaviour of a cluster spatio-temporal point pattern. However, the spatial aggregation decreases with increasing spatial distances, while the temporal aggregation is kept throughout most of the temporal range, as clearly shown in the right panel of Figure 2(a). This result is a consequence of many reports of IMD occurring close in space and time, and thus for short temporal periods it is quite likely that at least two reports of IMD occur close enough of each other. Additionally, the spatial aggregation shows the same behaviour even during periods of time sufficiently large. One way to emphasis this clustering behaviour is to compare the empirical surface of the product density function for IMD with the theoretical one for a Poisson point pattern with equal expected number of points than IMD. This result is clearly expected after visual inspection of Figure 3, and goes in the line found by Meyer et al. (2012)).

The left panel of Figure 4 shows the 95%-envelope surfaces obtained from 39 simulations of a spatio-temporal Poisson point pattern, see Møller and Ghorbani (2010) and Møller and Waagepetersen (2004)), together with the empirical product density. This figure shows how the empirical surface of the product density function for the IMD is larger than the upper 95%-envelope for small spatial and temporal distances.

The right panel of Figure 4 shows the confidence surface under a Poisson pattern based on the estimated $\widehat{\rho^{(2)}} \pm 2 \times$ standard deviations calculated using the close form of the variance in Section 2. We also superimpose the empirical product density for the IMD data. Again, the



(a) Average surface. (b) Average contour.

Figure 3: Product density kernel estimator for the spatio-temporal IMD dataset with $\epsilon = 13.9686$ km, and $\delta = 28$ days.

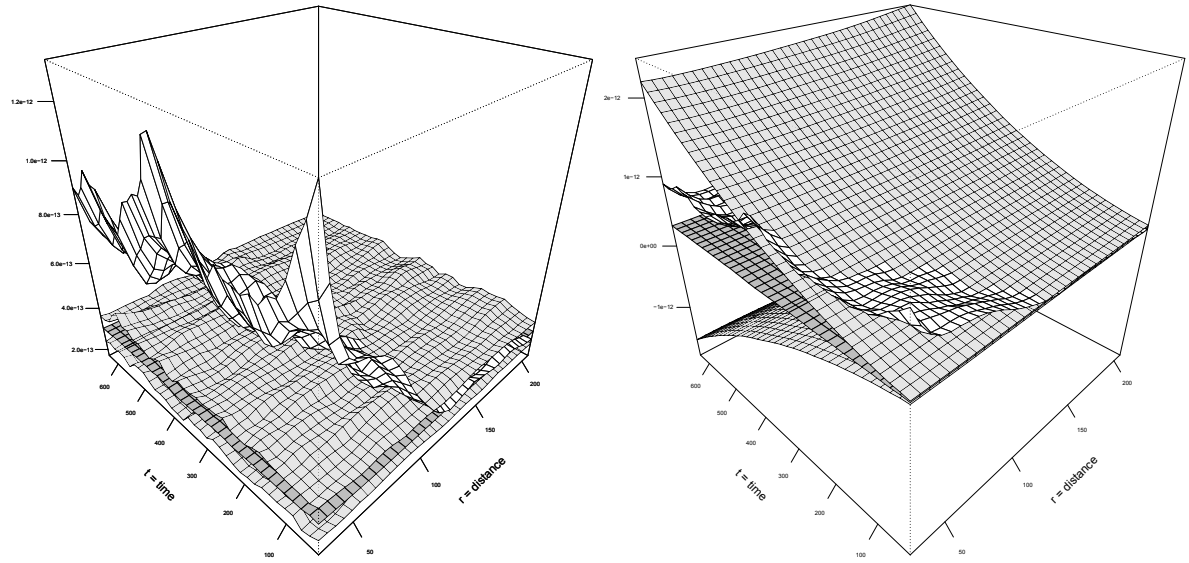
empirical density goes out the upper confidence surface. These two figures reveal that IMD has a contagious behaviour in their immediate spatio-temporal neighborhoods. These are solid arguments to reject the hypothesis of complete randomness in favour of a clustering structure.

6 Discussion

The spatio-temporal inhomogeneous product density function describes second-order characteristics of point processes. It is useful to analysis the spatio-temporal structure of the underlying point process, and thus provides a natural starting point for the analysis of spatio-temporal point process data. It can be considered an exploratory tool, for testing spatio-temporal clustering or spatio-temporal interaction.

We have proposed a non-parametric edge-corrected kernel estimate of the product density under the second-order intensity-reweighted stationary hypothesis. The expectation and variance of the estimator are obtained, and close expressions are derived under the Poisson case. First- and second-order spatio-temporal separability has also been considered and discussed. It is known (see Gabriel (2013)) that the performance of the pair correlation function and K-function can be severely altered by the intensity estimate. This can be explained by over-parametrisation or over-fitting in the case of a parametric estimation of the intensity function, or by the incapacity of distinguish first- and second-order effects from a single realisation of the point process in the case of a kernel-based estimation. This is in any case a kind of handicap and weakness in using these characteristics. We postulate the use of the product density as it provides the same amount of information, but with the added value that there is no need to estimate the intensity function.

We have provided sufficiently statistical grounds in favour of using this second-order tool in the practical analysis of spatio-temporal point patterns. However, we have based under developments on the hypothesis of second-order intensity-reweighted stationarity. The statistical properties of the spatio-temporal product density under general non-stationarity conditions or anisotropic structures



(a) 95%-envelope surfaces. (b) Confidence surfaces.

Figure 4: (a) Empirical surface for IMD dataset (white surface), and average and pointwise 95% envelope surfaces calculated from 39 simulations under Poisson process (grey surfaces). (b) Empirical surface for IMD dataset (white surface), and average and pointwise 95% confidence surfaces under Poisson process (grey surfaces).

remains an open problem.

Acknowledgements

Francisco J. Rodríguez-Cortés's research was supported by grant P1-1B2012-52. Mohammad Ghorbani's research was supported by the Center for Stochastic Geometry and Advanced Bioimaging, funded by a grant from the Villum Foundation. Jorge Mateu's research was supported by grant MTM2010-14961 from Ministery of Education.

References

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

Chiu, S. N., Stoyan, D., Kendall, W. S. and Mecke, J. (2013). *Stochastic Geometry and Its Applications*, third edn, John Wiley & Sons.

Cressie, N. and Collins, L. B. (2001a). Analysis of spatial point patterns using bundles of product density lisa functions, *Journal of Agricultural, Biological, and Environmental Statistics* **6**: 118135.

Cressie, N. and Collins, L. B. (2001b). Patterns in spatial point locations: Local indicators of spatial association in a minefield with clutter, *Naval Research Logistics* **48**: 333347.

Daley, D. J. and Vere-Jones, D. (2008). *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure*, second edn, Springer-Verlag, New York.

Diggle, P. J. (2006). Spatio-temporal point processes, partial likelihood, foot and mouth disease, *Statistical Methods in Medical Research* **15**: 325–336.

Diggle, P. J. (2013). *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*, Chapman and Hall/CRC, Boca Raton.

Fiksel, T. (1988). Edge-corrected density estimators for point processes, *Statistics* **19**: 67–75.

Gabriel, E. (2013). Estimating second-order characteristics of inhomogeneous spatio-temporal point processes: in uence of edge correction methods and intensity estimates, *Methodology and Computing in Applied Probability* .

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

Gabriel, E., Rowlingson, B. and Diggle, P. J. (2012). Stpp: Space-time point pattern simulation, visualisation and analysis. R package version 0.2.

Gabriel, E., Wilson, D. J., Leatherbarrow, A. J., Cheesbrough, J., Gee, S., Bolton, E., Fox, A., Fearnhead, P., Hart, C. A. and Diggle, P. J. (2010). Spatio-temporal epidemiology of campylobacter jejuni enteritis, in an area of northwest england, 20002002, *Epidemiology and infection* **138(10)**: 13841390.

Gelfand, A. E., Diggle, P. J., Guttorp, P. and Fuentes, M. (2010). *Handbook of Spatial Statistics*, CRC Press, Boca Raton.

Ghorbani, M. (2013). Testing the weak stationarity of a spatio-temporal point process, *Stochastic Environonmental Research and Risk Assessesment* **27**: 517–524.

Illian, J., Penttinen, A., Stoyan, H. and Stoyan, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*, John Wiley and Sons, Chichester.

Meyer, S., Elias, J. and M. Höhle, M. (2012). A spacetime conditional intensity model for invasive meningococcal disease occurrence, *Biometrics* **68(2)**: 607616.

Møller, J. and Diaz-Avalos, C. (2010). Structured spatio-temporal shot-noise Cox point process models, with a view to modelling forest fires, *Scandinavian Journal of Statistics* **37**: 2–25.

Møller, J. and Ghorbani, M. (2010). Second-order analysis of structured inhomogeneous spatio-temporal point processes, *Technical Report R-07-2010*, Department of Mathematical Sciences, Aalborg University. Submitted for journal publication.

Møller, J. and Ghorbani, M. (2012). Aspects of second-order analysis of structured inhomogeneous spatio-temporal point processes, *Statistica Neerlandica* **66**(4): 472–491.

Møller, J. and Ghorbani, M. (2013). Functional summary statistics for the johnson-mehl model, *Journal of Statistical Computation and Simulation* .

Møller, J. and Waagepetersen, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*, Chapman and Hall/CRC, Boca Raton.

Ohser, J. (1983). On estimators for the reduced second moment measure of point processes, *Mathematische Operationsforschung und Statistik, series Statistics* **14**: 63–71.

Ripley, B. D. (1988). *Statistical Inference for Spatial Processes*, Cambridge University Press, Cambridge.

Stoyan, D. (1987). Statistical analysis of spatial point processes: A soft-core model and cross-correlations of marks, *Biometrical Journal* **29**: 971–980.

Stoyan, D., Bertram, U. and Wendrock, H. (1993). Estimation variances for estimators of product densities and pair correlation functions of planar point processes, *Annals of the Institute of Statistical Mathematics* **45**: 211–221.

Stoyan, D., Kendall, W. S. and Mecke, J. (1995). *Stochastic Geometry and Its Applications*, second edn, Wiley, Chichester.

Stoyan, D. and Stoyan, H. (1994). *Fractals, Random Shapes and Point Fields*, Wiley, Chichester.

Wand, M. and Ripley, B. (2013). Functions for kernel smoothing for wand and jones (1995). R package version 2.23-10.

Appendix A :Moments of the product density estimator under spatio-temporal separability

For non-negative Borel functions h_1 and h_2 defined on $(\mathbb{R}^2)^{\otimes n}$ and $\mathbb{R}^{\otimes n}$ respectively, we assume that $h((\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)) = h_1(\mathbf{u}_1, \ldots, \mathbf{u}_n)h_2(s_1, \ldots, s_n)$, and considering n-order spatio-temporal separability we can rewrite (1) as

$$\begin{aligned}\mathbb{E}\sum_{\Theta_n}^{\neq} h((\mathbf{u}_1, s_1), \ldots, (\mathbf{u}_n, s_n)) &= \int_{W^{\otimes n}} h_1(\mathbf{u}_1, \ldots, \mathbf{u}_n)\bar{\rho}_1^{(n)}(\mathbf{u}_1, \ldots, \mathbf{u}_n) \prod_{i=1}^{n} \mathrm{d}\mathbf{u}_i \\ &\quad \times \int_{T^{\otimes n}} h_2(s_1, \ldots, s_n)\bar{\rho}_2^{(n)}(s_1, \ldots, s_n) \prod_{i=1}^{n} \mathrm{d}s_i \\ &= \mathbb{E} \sum_{\mathbf{u}_1,\ldots,\mathbf{u}_n \in X_{\text{space}}}^{\neq} f_1(\mathbf{u}_1, \ldots, \mathbf{u}_n)\mathbb{E} \sum_{s_1,\ldots,s_n \in X_{\text{time}}}^{\neq} f_2(s_1, \ldots, s_n),\end{aligned} \tag{20}$$

where

$$f_1(\mathbf{u}_1, \ldots, \mathbf{u}_n) = h_1(\mathbf{u}_1, \ldots, \mathbf{u}_n)/\left(\int_{W^{\otimes n}} \bar{\rho}_1^{(n)}(\mathbf{u}_1, \ldots, \mathbf{u}_n) \prod_{i=1}^{n} \mathrm{d}\mathbf{u}_i\right)$$

and

$$f_2(s_1, \ldots, s_n) = h_2(s_1, \ldots, s_n)/\left(\int_{T^{\otimes n}} \bar{\rho}_2^{(n)}(s_1, \ldots, s_n) \prod_{i=1}^{n} \mathrm{d}s_i\right)$$

.

A.1: Expectation

Combining (15) and (20) for $n = 2$ we have,

$$\mathbb{E}\left[\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right] = \frac{(n-2)!}{n!}\mathbb{E}\sum_{\mathbf{u}_i,\mathbf{u}_j\in X_{\text{space}}}^{\neq}\frac{\kappa_{1\epsilon}(\|\mathbf{u}_i-\mathbf{u}_j\|-r)}{2\pi\gamma_W(r)r}\mathbb{E}\sum_{s_i,s_j\in X_{\text{time}}}^{\neq}\frac{\kappa_{2\delta}(|s_i-s_j|-t)}{2\gamma_T(t)}$$

$$= \int_{-r/\epsilon}^{\infty}\int_{-t/\delta}^{\infty}\frac{\kappa_1(u)\kappa_2(v)\gamma_W(r+\epsilon u)\gamma_T(t+\delta v)}{r\gamma_W(r)\gamma_T(t)}\rho^{(2)}(r+\epsilon u, t+\delta v)(r+\epsilon u)\mathrm{d}u\,\mathrm{d}v$$

Note that under separability we obtain the same expression (16) as in the general case.

A.2: Second-order moment

Under the same assumptions as in the above case for $n = 3, 4$ in (20), and using (14) the second-order moment of the product density estimator under separability is given by

$$\mathbb{E}\left[\left(\widehat{\rho^{(2)}}_{\epsilon,\delta}(r,t)\right)^2\right] = \frac{(c(r,t))^2}{16}\Big[2E_1(B) + 4E_2(B) + E_3(B)\Big],$$

where $E_1(B)$, $E_2(B)$ and $E_3(B)$ are slightly different from their general corresponding and can be obtained by the same procedure as expectation under separability. So, expression (17) as in the general case is obtained. Thus all results and properties for the general case are also satisfied under the separable case.