Spatial Statistics 9 (2014) 20–37

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

On the local odds ratio between points and marks in marked point processes

Tonglin Zhang [a,*], Qianlai Zhuang [b,c]

[a] Department of Statistics, Purdue University, 250 North University Street, West Lafayette, IN 47907-2066, USA
[b] Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, West Lafayette, IN, 47907, USA
[c] Department of Agronomy, Purdue University, West Lafayette, IN, 47907, USA

ARTICLE INFO

Article history:
Received 26 August 2013
Accepted 9 December 2013
Available online 24 December 2013

Keywords:
Consistency
Contingency table
Intensity functions
Local odds ratios
Marked point processes
Relative risks

ABSTRACT

Marked point processes are widely used stochastic models for representing a finite number of natural hazard events located in space and time and their data often associate event measurements (i.e. marks) with event locations (i.e. points). An interesting statistical problem of marked point processes is to measure and estimate the localized dependence between points and marks. To solve this problem, an approach of local odds ratio is proposed, where the local odds ratio is defined by the localized ratio of the relative risk for an event to have a small mark to the relative risk to have a large mark. To establish the approach, the article presents definition, estimation, and statistical properties. To justify the usefulness of the approach, the article presents two particular examples in natural hazards: a forest wildfire study and an earthquake study. It finds that values of local odds ratios are mostly likely low in one subarea but high in another subarea, which indicates that events with large mark values are mostly likely to appear in the former subarea but less likely to appear in the latter subarea. It is expected that the proposed approach will be widely applicable in natural hazard studies.

1. Introduction

Marked point processes (MPPs) are commonly used stochastic models, which have been widely applied to data involved both the spatial (including spatiotemporal) coordinates of events and their corresponding measurements. Methods of MPPs are often used to model a number of natural hazards

* Corresponding author. Tel.: +1 765 496 2096; fax: +1 765 494 0558.
E-mail addresses: tlzhang@purdue.edu, tlzhang@stat.purdue.edu (T. Zhang), qzhuang@purdue.edu (Q. Zhuang).

2211-6753/$ – see front matter
http://dx.doi.org/10.1016/j.spasta.2013.12.002

located in space and time. There are many successful applications of MPPs in literature. These include MPP modeling and prediction of earthquakes (Holden et al., 2003; Ogata, 1988; Ogata and Katsura, 1993; Vere-Jones, 1995), where each earthquake is represented by a magnitude and a space–time coordinate. The three-dimensional space coordinate contains the longitude, latitude and depth of earthquake occurrences. MPPs for forest wildfires have been discussed by Peng et al. (2005), where each wildfire is represented by its area burned and space–time coordinate. The two-dimensional space coordinate contains the longitude and latitude of wildfire occurrences.

The approach of the localized dependence between points and marks relies on a strict mathematical definition of MPPs, which will be introduced in the next section. In a simplified version, one can describe an MPP with an artificial order such that data can be represented by $N = \{(S_i, M_i) : i = 1, \ldots, n\}$, where n is the total number of events, S_i are the locations of points, and M_i are the corresponding marks. Specific geostatistical methods including variogram analysis, various kinds of kriging, and geostatistical simulation techniques may be used to model an MPP (Cressie, 1993), but these methods rely on the fundamental assumption that point locations appear independently of marks because the definition of correlation functions used in a geostatistical method often ignores point distributions (Diggle et al., 2003). Since the independence assumption between points and marks is often unrealistic in applications, these methods may not be used if points and marks are highly correlated. For instance, the relative positions of trees in a forest have repercussions on their size owing to their competition for light or nutrient (Schlather et al., 2004), indicating that tree sizes and locations of trees may not be independent. Forest wildfire activities exhibit power-law relationships between frequency and burned area (Malamud et al., 2005), indicating that the burned area and the locations of forest wildfires may not be independent either.

It is especially convenient in modeling, estimation, and prediction in an MPP if marks and points are independent. Many commonly used Hawkes models, such as the epidemic-type aftershock sequences (EATS) model (Ogata, 1998), may exhibit the independence between marks and points (Schoenberg, 2004). In the **spatstat** (Baddeley and Turner, 2005) and **PtProcess** (Harte, 2010) packages in **R** several useful methods based on MPPs under the assumption of independence are available (McElroy and Politis, 2007; Poliltis and Sherman, 2001). If the independence assumption is violated, then intensity-dependent models may also be useful (Ho and Stoyan, 2008; Malinowski et al., 2012; Myllymäki and Penttinen, 2009). However, these methods cannot be used to describe the localized dependence between points and marks because the relationship between points and marks is often modeled globally. For example, the mark (magnitude) distribution in earthquake activities locally depends on their geographical locations which cannot be accounted for by an intensity-dependent model. The mark (area burned) distribution in forest wildfire activities locally depends on forest densities which cannot either be accounted for by an intensity-dependent model. Therefore, it is important to develop a statistical approach to modeling the local dependence between points and marks.

To account for the dependence between marks and points, we modify the approach of the odds ratio for contingency tables (Agresti, 2002) to an approach of local dependence between points and marks of MPPs. We call it the approach of the local odds ratio, where the local odds ratio is a measure of the strength of the local dependence between points and marks. In the approach, we note that the odds ratio is one of the most important measures of row–column dependence in a contingency table and it is also an important index in binomial or Poisson regression. Unlike other measures of dependence in the contingency table (such as the relative risk), the odds ratio treats rows and columns symmetrically. Its value does not change when the orientation of the table reverses so that the rows become the columns and the columns become the rows. Therefore, the odds ratio is invariant under the transpose transformation. To define the localized odds ratio, we modify the classical definition of the odds ratio. We expect that the modified local odds ratio can be theoretically derived at any given location in the whole study area. Based on values of the local odds ratio, one can compare the local risks with the global risks for large mark events. In addition, one can compare local risks between two specific locations. Values of local odds ratio are useful to identify a subarea with high risks of large events. Examples include a method to identify a subarea of high risk of large earthquakes in earthquake studies or a subarea of high risks of large fires in forest wildfire studies. Because large events are more important than small events in the natural hazard studies, the local odds ratio may be used as a standard measure for the risk analysis of large events in these studies.

The paper is organized as follows. Section 2 reviews the concept of MPPs and provides the statistical definition of the local odds ratio. Section 3 provides estimation of the local odds ratio as well as a proof of the consistency of the estimator. Section 4 provides a simulation study to evaluate the performance of the estimator. Section 5 applies the approach to a forest wildfire dataset and an earthquake dataset. Section 6 provides a discussion.

2. Method

To establish the approach, we briefly review the definition of MPPs in spatial statistics and the definition of odds ratio in contingency tables. Based on the reviews, we provide the definition of the local odds ratio. Existence and uniqueness of the local odds ratio in the definition as well as its statistical properties are examined.

2.1. Definitions of MPPs

The definition of MPPs is well-established and can be found in many textbooks (e.g. in Daley and Vere-Jones (2003) and Karr (1991)). Overall, an MPP is a pure point process defined on the product space of points and marks, but the concept has its own life in applications. Let $\mathcal{S}$ and $\mathcal{M}$ be complete separable metric spaces. Let $\mathscr{S}$ and $\mathscr{M}$ be the collections of Borel sets in $\mathcal{S}$ and $\mathcal{M}$, respectively. An MPP N with points in $\mathcal{S}$ and marks in $\mathcal{M}$ is a point process on $\mathcal{S} \times \mathcal{M}$ with the additional property that the underlying point process N_s is itself a point process and for any bounded $A \in \mathscr{S}$ there is $N_s(A) = N(A \times \mathcal{M}) < \infty$, where $N(A \times B)$ is the number of points in $A \times B$ for $A \in \mathscr{S}$ and $B \in \mathscr{M}$, respectively. Denote $n = N(\mathcal{S} \times \mathcal{M})$ and assume n is finite. Then, n is a discrete random variable. In modeling the occurrence of ecological or geological events when depth is not involved, we may have $\mathcal{S} = R^d$ with $d = 2$ if time is not considered or $d = 3$ if time is considered. In addition, three-dimensional point patterns may also occur in space (Stein et al., 2011) and in this case we have $d = 3$ if time is not considered or $d = 4$ if time is considered. It has been pointed out that the distribution of N can be uniquely determined by the joint distribution of $\{N(A_i \times B_j) : i = 1, \ldots, I; j = 1, \ldots, J\}$ for any partition $\{A_1, \ldots, A_I\} \in \mathscr{S}$ of $\mathcal{S}$ and $\{B_1, \ldots, B_J\} \in \mathscr{M}$ of $\mathcal{M}$. Based on the distribution of N, we can define the kth order intensity function of N (if it exists) as

$$\lambda_k[(\mathbf{s}_1, m_1), \ldots, (\mathbf{s}_k, m_k)] = \lim_{|d\mathbf{s}_i \times dm_i| \to 0, i=1,\ldots,k} \left\{ \frac{E[N(d\mathbf{s}_1 \times dm_1) \cdots N(d\mathbf{s}_k \times dm_k)]}{|d\mathbf{s}_1 \times dm_1| \cdots |d\mathbf{s}_k \times dm_k|} \right\}, \tag{1}$$

where $(\mathbf{s}_i, m_i)$ are distinct pairs of points and marks in $\mathcal{S} \times \mathcal{M}$, $d\mathbf{s}_i \times dm_i$ is an infinitesimal region containing $(\mathbf{s}_i, m_i) \in \mathcal{S} \times \mathcal{M}$, and $|d\mathbf{s}_i \times dm_i|$ is the Lebesgue measure of $d\mathbf{s}_i \times dm_i$.

For convenience, we denote $\lambda(\mathbf{s}, m) = \lambda_1[(\mathbf{s}, m)]$ and always assume that the k-th order intensity functions of N exists for $k \leq 4$ in the rest of the paper. With this assumption, the moments up to four of $N(A \times B)$ exists. The formulae of the moments of MPPs can be easily derived by formulae of pure spatial point process, which are available in many articles (e.g. in Moller and Waagepetersen (2007)). Here, we only display some of those that are useful in the paper.

Let the mean measure of N be denoted by

$$\mu(A \times B) = \int_A \int_B \lambda(\mathbf{s}, m) dm d\mathbf{s}$$

and the second-order moment measure be denoted by

$$\mu^{(2)}(A_1 \times B_1, A_2 \times B_2) = \int_{A_1} \int_{B_1} \int_{A_2} \int_{B_2} \lambda_2[(\mathbf{s}_1, m_1), (\mathbf{s}_2, m_2)] dm_2 d\mathbf{s}_2 dm_1 d\mathbf{s}_1.$$

Then, the mean structure of N is

$$E[N(A \times B)] = \mu(A \times B) \tag{2}$$

Table 1
Notation for probabilities in a 2 × 2 contingency table.

Row	Column		Total
	1	2	
1	π_{11}	π_{12}	π_{1+}
2	π_{21}	π_{22}	π_{2+}
Total	π_{+1}	π_{+2}	1

and the covariance structure of N is

$$\begin{aligned} &Cov[N(A_1 \times B_1), N(A_2 \times B_2)] \\ &= \mu^{(2)}(A_1 \times B_1, A_2 \times B_2) + \mu[(A_1 \cap A_2) \times (B_1 \cap B_2)] - \mu(A_1 \times B_1)\mu(A_2 \times B_2) \\ &= \int_{A_1}\int_{B_1}\int_{A_2}\int_{B_2} \{\lambda_2[(\mathbf{s}_1, m_1), (\mathbf{s}_2, m_2)] - \lambda(\mathbf{s}_1, m_1)\lambda(\mathbf{s}_2, m_2)\} dm_2 d\mathbf{s}_2 dm_1 d\mathbf{s}_1 \\ &\quad + \mu[(A_1 \cap A_2) \times (B_1 \cap B_2)]. \end{aligned} \tag{3}$$

In order to characterize the dependence between points and marks in the product space after adjusting the effect of the first-order intensity function, it is useful to consider the pair correlation function

$$g[(\mathbf{s}_1, m_1), (\mathbf{s}_2, m_2)] = \frac{\lambda_2[(\mathbf{s}_1, m_1), (\mathbf{s}_2, m_2)]}{\lambda(\mathbf{s}_1, m_1)\lambda(\mathbf{s}_2, m_2)} \tag{4}$$

if $\lambda(\mathbf{s}_1, m_1)$ and $\lambda(\mathbf{s}_2, m_2)$ are positive. If N is a marked Poisson process, then $g[(\mathbf{s}_1, m_1), (\mathbf{s}_2, m_2)]$ is always equal to one at any $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{S}$ and $m_1, m_2 \in \mathcal{M}$. Otherwise, N either contains attractions or repulsions among events. Based on the pair correlation function, the covariance structure of N becomes

$$\begin{aligned} &Cov[N(A_1 \times B_1), N(A_2 \times B_2)] \\ &= \int_{A_1}\int_{B_1}\int_{A_2}\int_{B_2} \{g[(\mathbf{s}_1, m_1), (\mathbf{s}_2, m_2)] - 1\}\lambda(\mathbf{s}_1, m_1)\lambda(\mathbf{s}_2, m_2) dm_2 d\mathbf{s}_2 dm_1 d\mathbf{s}_1 \\ &\quad + \mu[(A_1 \cap A_2) \times (B_1 \cap B_2)]. \end{aligned} \tag{5}$$

2.2. Reviews of odds ratio

The classical definition of odds ratio can be found in many textbooks (e.g. page 44 in Agresti (2002)). The definition is usually based on a 2 × 2 contingency table. Suppose the 2 × 2 contingency table is expressed by rows and columns as the one displayed in Table 1, where π_{ij} is the probability for an event to fall in cell (i, j), π_{i+} is the probability for an event to fall in the ith row, and π_{+j} is the probability for an event to fall in the jth column, where $i, j = 1, 2$. Note that $\pi_{1|1} = \pi_{11}/\pi_{1+}$ is the conditional probability for an event to fall in the first column conditioning on the event to fall in the first row, and $\pi_{1|2} = \pi_{21}/\pi_{2+}$ is the conditional probability for an event to fall in the first column conditioning on the event to fall in the second row. The relative risk is defined as

$$r_1 = \frac{\pi_{1|1}}{\pi_{1|2}} = \frac{\pi_{11}(\pi_{21} + \pi_{22})}{\pi_{21}(\pi_{11} + \pi_{12})}. \tag{6}$$

Similarly, the relative risk for the second column is $r_2 = [\pi_{12}(\pi_{21} + \pi_{22})]/[\pi_{22}(\pi_{11} + \pi_{12})]$.

In Table 1, the odds for the first row is defined as $\Omega_1 = \pi_{11}/\pi_{12}$ and the odds for the second row is defined as $\Omega_2 = \pi_{21}/\pi_{22}$. The odds ratio is defined by the ratios of Ω_1 and Ω_2 as

$$\theta = \frac{\Omega_1}{\Omega_2} = \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}} = \frac{r_1}{r_2}. \tag{7}$$

The odds ratio can equal any nonnegative number. The condition $\theta = 1$ corresponds to independence of rows and columns. If $1 < \theta < \infty$, the first column in row one is more likely to occur than

the first column in row two; otherwise the first column in row one is less likely to occur than the first column in row two. Values of θ farther from 1.0 in a direction represent stronger dependence. If one value is the reverse of the other, then they represent the same strength of dependence but in opposite directions.

Because the probability table is not available in practice, the sample odds ratio is used to describe the dependence between rows and columns in Table 1. Let n_{ij} be the observed count in cell (i, j). Then, the sample odds ratio is

$$\hat{\theta} = \frac{n_{11}n_{22}}{n_{12}n_{21}}.$$

The sample odds ratio is an estimator of odds ratio, which goes to θ in probability as the sample size goes to infinity.

2.3. Definition of local odds ratio

We use the idea of the odds ratio to define the local dependence between points and marks. To construct a 2×2 contingency table, we compare mark values greater than or equal to m_0 with mark values less than m_0, where m_0 is a pre-selected threshold which relies on the particular interest in applications. For example, an earthquake with magnitude greater than or equal to 6.0 is considered as a large earthquake and a forest wildfire with area burned greater than 2 km^2 is considered as a large fire. We may choose $m_0 = 6$ in earthquake studies and $m_0 = 2$ km^2 in forest wildfire studies, which indicates that the relative risk of large earthquakes or large forest wildfires is compared with the relative risks of small earthquakes or small forest wildfires.

Suppose the domain of marks is an open sub-interval of real numbers, which is denoted by $\mathcal{M} = (\underline{m}, \overline{m})$. Let $U_{\mathbf{s}}$ be a neighborhood of $\mathbf{s} \in \mathcal{S}$ and $\overline{U}_{\mathbf{s}}$ be its complementary set. Denote $d(U_{\mathbf{s}})$ as the diameter of $U_{\mathbf{s}}$, which is defined by $d(U_{\mathbf{s}}) = \max\{\rho(\mathbf{s}, \mathbf{s}') : \mathbf{s}, \mathbf{s}' \in U_{\mathbf{s}}\}$ for a distance ρ. For a pre-selected $m_0 \in (\underline{m}, \overline{m})$, let $n_{11} = N(\{\mathbf{s} \in U_{\mathbf{s}}, m < m_0\})$, $n_{12} = N(\{\mathbf{s} \in U_{\mathbf{s}}, m \geq m_0\})$, $n_{21} = N(\{\mathbf{s} \notin U_{\mathbf{s}}, m < m_0\})$, and $n_{22} = N(\{\mathbf{s} \notin U_{\mathbf{s}}, m \geq m_0\})$. Then, $E(n_{11}) = \int_{U_{\mathbf{s}}} \int_{\underline{m}}^{m_0} \lambda(\mathbf{s}, m) dm d\mathbf{s}$, $E(n_{12}) = \int_{U_{\mathbf{s}}} \int_{m_0}^{\overline{m}} \lambda(\mathbf{s}, m) dm d\mathbf{s}$, $E(n_{21}) = \int_{\overline{U}_{\mathbf{s}}} \int_{\underline{m}}^{m_0} \lambda(\mathbf{s}, m) dm d\mathbf{s}$, and $E(n_{22}) = \int_{\overline{U}_{\mathbf{s}}} \int_{m_0}^{\overline{m}} \lambda(\mathbf{s}, m) dm d\mathbf{s}$. Denote $\pi_{ij} = E(n_{ij})/E(n)$. Then, π_{11}, π_{12}, π_{21}, and π_{22} can be used to define an contingency table as we have displayed in Table 1. An odds ratio is derived as

$$\begin{aligned}\theta_{m_0}(U_{\mathbf{s}}) &= \frac{\pi_{11}\pi_{22}}{\pi_{12}\pi_{21}} = \frac{E(n_{11})E(n_{22})}{E(n_{21})E(n_{12})} \\ &= \frac{\left[\int_{U_{\mathbf{s}}} \int_{\underline{m}}^{m_0} \lambda(\mathbf{s}, m) dm d\mathbf{s}\right]\left[\int_{\overline{U}_{\mathbf{s}}} \int_{m_0}^{\overline{m}} \lambda(\mathbf{s}, m) dm d\mathbf{s}\right]}{\left[\int_{U_{\mathbf{s}}} \int_{m_0}^{\overline{m}} \lambda(\mathbf{s}, m) dm d\mathbf{s}\right]\left[\int_{\overline{U}_{\mathbf{s}}} \int_{\underline{m}}^{m_0} \lambda(\mathbf{s}, m) dm d\mathbf{s}\right]}. \end{aligned} \tag{8}$$

The local odds ratio between marks below and above m_0 at $\mathbf{s}$ is defined as

$$\theta_{m_0}(\mathbf{s}) = \lim_{d(U_{\mathbf{s}}) \to 0} \theta_{m_0}(U_{\mathbf{s}}) = \frac{\psi_{11}(\mathbf{s}, m_0)\psi_{22}(m_0)}{\psi_{12}(\mathbf{s}, m_0)\psi_{21}(m_0)}, \tag{9}$$

where $\psi_{11}(\mathbf{s}, m_0) = \int_{\underline{m}}^{m_0} \lambda(\mathbf{s}, m) dm$, $\psi_{12}(\mathbf{s}, m_0) = \int_{m_0}^{\overline{m}} \lambda(\mathbf{s}, m) dm$, $\psi_{21}(m_0) = \int_{\mathcal{S}} \int_{\underline{m}}^{m_0} \lambda(\mathbf{s}, m) dm d\mathbf{s}$, and $\psi_{22}(m_0) = \int_{\mathcal{S}} \int_{m_0}^{\overline{m}} \lambda(\mathbf{s}, m) dm d\mathbf{s}$. It is clear that the value of $\theta_{m_0}(\mathbf{s})$ depends on both m_0 and $\mathbf{s}$.

Theorem 1. *The local odds ratio $\theta_{m_0}(\mathbf{s})$ is well-defined and is positive and continuous in all $\mathbf{s} \in \mathcal{S}$ and $m_0 \in \mathcal{M}$ if $\lambda(\mathbf{s}, m)$ satisfies all of the following conditions:*

(A1) $\lambda(\mathbf{s}, m) > 0$ *for all* $\mathbf{s} \in \mathcal{S}$ *and* $m \in \mathcal{M}$;
(A2) $\lambda(\mathbf{s}, m)$ *is continuous in all* $\mathbf{s} \in \mathcal{S}$ *and* $m \in \mathcal{M}$;
(A3) $\int_{\mathcal{S}} \int_{\mathcal{M}} \lambda(\mathbf{s}, m) dm d\mathbf{s} < \infty$.

Proof. Conclusion can be directly drawn because under the assumption of the theorem we have

$$\lim_{d(U_{\mathbf{s}})\to 0}\frac{1}{|U_{\mathbf{s}}|}\int_{U_{\mathbf{s}}}\int_{B}\lambda(\mathbf{s},m)dmd\mathbf{s}=\int_{B}\lambda(\mathbf{s},m)dm$$

and

$$\lim_{d(U_{\mathbf{s}})\to 0}\int_{\overline{U}_{\mathbf{s}}}\int_{B}\lambda(\mathbf{s},m)dmd\mathbf{s}=\int_{\mathcal{S}}\int_{B}\lambda(\mathbf{s},m)dmd\mathbf{s}$$

and both are positive, finite, and continuous in $\mathbf{s}$ and m_0. Choosing $B=(\underline{m},m_0)$ and $\overline{B}=[m_0,\overline{m})$ above and applying them to Eq. (8), we have the conclusion of the theorem. ◇

The local odds ratio has good interpretations. According to the definition given by (9), it is clear that

$$\frac{\psi_{21}(m_0)}{\psi_{22}(m_0)}=\frac{\int_{\mathcal{S}}\int_{\underline{m}}^{m_0}\lambda(\mathbf{s},m)dmd\mathbf{s}}{\int_{\mathcal{S}}\int_{m_0}^{\overline{m}}\lambda(\mathbf{s},m)dmd\mathbf{s}}$$

represents the odds between small marks and large marks in the whole study area. The ratio

$$\frac{\psi_{11}(\mathbf{s},m_0)}{\psi_{12}(\mathbf{s},m_0)}=\frac{\int_{\underline{m}}^{m_0}\lambda(\mathbf{s},m)dm}{\int_{m_0}^{\overline{m}}\lambda(\mathbf{s},m)dm}$$

represents the localized odds between small marks and large marks at $\mathbf{s}$. The ratio

$$\frac{\psi_{11}(\mathbf{s},m_0)}{\psi_{21}(m_0)}=\frac{\int_{\underline{m}}^{m_0}\lambda(\mathbf{s},m)dm}{\int_{\mathcal{S}}\int_{\underline{m}}^{m_0}\lambda(\mathbf{s},m)dmd\mathbf{s}}$$

represents the relative risk between $\mathbf{s}$ and the global average for small marks. The ratio

$$\frac{\psi_{12}(\mathbf{s},m_0)}{\psi_{22}(m_0)}=\frac{\int_{m_0}^{\overline{m}}\lambda(\mathbf{s},m)dm}{\int_{\mathcal{S}}\int_{m_0}^{\overline{m}}\lambda(\mathbf{s},m)dmd\mathbf{s}}$$

represents the relative risk between $\mathbf{s}$ and the global average for large marks. Thus, $\theta_{m_0}(\mathbf{s})$ is either the ratio of local odds to the global odds or the ratio of the relative risks between small events and large events. If $\theta_{m_0}(\mathbf{s})$ is small, then large events are more likely to occur around $\mathbf{s}$. The value of $\theta_{\mathbf{s}}(m_0)>1$ for a particular $\mathbf{s}$ corresponds to the case that small events are more likely to appear around $\mathbf{s}$ than the global level. The value of $\theta_{\mathbf{s}}(m_0)<1$ corresponds to the case that large events are more likely to appear around $\mathbf{s}$ than the global level. In addition, the value of $\theta_{m_0}(\mathbf{s})/\theta_{m_0}(\mathbf{s}')$ can be used to compare the relative risks for large events between two specific locations. If the value is less than one, large events are more likely to occur around $\mathbf{s}$ than $\mathbf{s}'$; otherwise, small events are more likely to occur. Therefore, values of $\theta_{m_0}(\mathbf{s})$ in the whole study area $\mathcal{S}$ can be used to describe how likely to see a small (or a large) event at the local level conditioning on the occurrence of an event at that site.

Theorem 2. *Assume Conditions* (A1)–(A3) *in Theorem* 1 *hold. A necessary condition for points and marks to be independent is* $\theta_{m_0}(\mathbf{s})=1$ *for all* $\mathbf{s}\in\mathcal{S}$ *and* $m_0\in\mathcal{M}$. *If* N *is a marked Poisson process, then the condition is also sufficient.*

Proof. According to Schoenberg (2004), there exists $f_1(\mathbf{s})$ and $f_2(m)$ such that $\lambda(\mathbf{s},m)=f_1(\mathbf{s})f_2(m)$ if points and marks are independent. Therefore, the condition that $\theta_{m_0}(\mathbf{s})=1$ for all $\mathbf{s}\in\mathcal{S}$ and $m_0\in\mathcal{M}$ is a necessary condition for points and marks to be independent. Now consider the case that N is a marked Poisson process with $\theta_{m_0}(\mathbf{s})=1$ for all $\mathbf{s}\in\mathcal{S}$ and $m_0\in\mathcal{M}$. Let

$$h_1(m_0)=\frac{\int_{\mathcal{S}}\int_{\underline{m}}^{m_0}\lambda(\mathbf{s},m)dmd\mathbf{s}}{\int_{\mathcal{S}}\int_{\underline{m}}^{\overline{m}}\lambda(\mathbf{s},m)dmd\mathbf{s}}$$

and

$$h_2(\mathbf{s}) = \int_{\underline{m}}^{\overline{m}} \lambda(\mathbf{s}, m)dm.$$

Then $h_1(m_0)$ is only a function of m_0 and $h_2(\mathbf{s})$ is only a function of $\mathbf{s}$. Therefore

$$\begin{aligned}\frac{\int_{m_0}^{\overline{m}} \lambda(\mathbf{s}, m)dm}{\int_{\underline{m}}^{m_0} \lambda(\mathbf{s}, m)dm} = \frac{\int_{\mathcal{S}}\int_{m_0}^{\overline{m}} \lambda(\mathbf{s}, m)dmd\mathbf{s}}{\int_{\mathcal{S}}\int_{\underline{m}}^{m_0} \lambda(\mathbf{s}, m)dmd\mathbf{s}} &\Rightarrow \frac{\int_{m_0}^{\overline{m}} \lambda(\mathbf{s}, m)dm}{\int_{\underline{m}}^{m_0} \lambda(\mathbf{s}, m)dm} + 1 = \frac{\int_{\mathcal{S}}\int_{m_0}^{\overline{m}} \lambda(\mathbf{s}, m)dmd\mathbf{s}}{\int_{\mathcal{S}}\int_{\underline{m}}^{m_0} \lambda(\mathbf{s}, m)dmd\mathbf{s}} + 1\\ &\Rightarrow \frac{\int_{\underline{m}}^{\overline{m}} \lambda(\mathbf{s}, m)dm}{\int_{\underline{m}}^{m_0} \lambda(\mathbf{s}, m)dm} = \frac{\int_{\mathcal{S}}\int_{\underline{m}}^{\overline{m}} \lambda(\mathbf{s}, m)dmd\mathbf{s}}{\int_{\mathcal{S}}\int_{\underline{m}}^{m_0} \lambda(\mathbf{s}, m)dmd\mathbf{s}}\\ &\Rightarrow \int_{\underline{m}}^{m_0} \lambda(\mathbf{s}, m)dm = h_1(m_0)h_2(\mathbf{s}).\end{aligned}$$

Differentiate the above equation on both sides with respect to m_0. Then, we have

$$\lambda(\mathbf{s}, m_0) = h_1'(m_0)h_2(\mathbf{s}),$$

which implies that points and marks are independent if N is a marked Poisson process (Schoenberg, 2004). ◇

Remark. Because $\lambda(\mathbf{s}, m) = f_1(\mathbf{s})f_2(m)$ is only a necessary condition of independence, we cannot conclude points and marks are independent if $\theta_{m_0}(\mathbf{s}) = 1$ for all $\mathbf{s} \in \mathcal{S}$ and $m_0 \in \mathcal{M}$ in a general MPP. However, if one can show $\lambda(\mathbf{s}, m)$ cannot always be written into $f_1(\mathbf{s})f_2(m)$, then it is enough to conclude points and marks are not independent. Recently, a few methods have been proposed to assess the independence between points and marks. These include tests for stationarity and isotropy of MPPs using variograms (Schlather et al., 2004), a nonparametric kernel-based test to assess the separability of the first-order intensity function (Schoenberg, 2004), and a χ^2-based test to assess the interaction between points and marks in N when N_s is stationary and isotropic (Guan and Afshartous, 2007). Using these methods, we are able to test whether $\theta_{m_0}(\mathbf{s}) = 1$ for all $\mathbf{s} \in \mathcal{S}$ and $m_0 \in \mathcal{M}$.

3. Estimation

We propose a nonparametric method to estimate $\theta_c(\mathbf{s})$. It is modified from the well-known kernel density estimation method which has been extensively used in nonparametric statistics. To apply the kernel method, we have to choose a kernel function $K(\mathbf{s})$ on R^d and a pre-selected bandwidth $h > 0$, where the kernel function satisfies

$$\int_{R^d} K(\mathbf{s})d\mathbf{s} = 1.$$

As long as $K(\mathbf{s})$ and h are decided, we need to estimate $\psi_{11}(\mathbf{s}, m_0)$, $\psi_{12}(\mathbf{s}, m_0)$, $\psi_{21}(m_0)$, and $\psi_{22}(m_0)$ in Eq. (9). We propose a two-step method to estimate these functions. In the first step, we estimate $\psi_{11}(\mathbf{s}, m_0)$ and $\psi_{12}(\mathbf{s}, m_0)$. We note that they represent intensity functions for points at $\mathbf{s}$ according to $m < m_0$ and $m \geq m_0$, respectively. Therefore, we have the estimator of $\psi_{11}(\mathbf{s}, m_0)$ as

$$\hat{\psi}_{11}(\mathbf{s}, m_0) = \frac{1}{h^d}\sum_{i=1}^{n} K\left(\frac{\mathbf{s} - \mathbf{s}_i}{h}\right) I(m_i < m_0)$$

and the estimator of $\psi_{12}(\mathbf{s}, m_0)$ as

$$\hat{\psi}_{12}(\mathbf{s}, m_0) = \frac{1}{h^d}\sum_{i=1}^{n} K\left(\frac{\mathbf{s} - \mathbf{s}_i}{h}\right) I(m_i \geq m_0).$$

In the second step, we estimate $\psi_{21}(m_0)$ and $\psi_{22}(m_0)$. We note that $\psi_{21}(m_0)$ and $\psi_{22}(m_0)$ represent cumulative measures for points in the whole study area according to $m < m_0$ and $m \geq m_0$,

respectively. Although they do not depend on $\mathbf{s}$, we consider the impact of $\hat{\psi}_{11}(\mathbf{s}, m_0)$ and $\hat{\psi}_{12}(\mathbf{s}, m_0)$ on estimation of $\theta_{m_0}(\mathbf{s})$ and propose the estimator of $\psi_{21}(m_0)$ as

$$\hat{\psi}_{21}(m_0) = \sum_{i=1}^{n} I(m_i < m_0) - \sum_{i=1}^{n} K\left(\frac{\mathbf{s} - \mathbf{s}_i}{h}\right) I(m_i < m_0) \tag{10}$$

and the estimator of $\psi_{22}(m_0)$ as

$$\hat{\psi}_{22}(m_0) = \sum_{i=1}^{n} I(m_i \geq m_0) - \sum_{i=1}^{n} K\left(\frac{\mathbf{s} - \mathbf{s}_i}{h}\right) I(m_i \geq m_0). \tag{11}$$

We adopt the above two estimators because the second term on the right side of (10) or (11) is used in estimation of $\psi_{11}(\mathbf{s}, m_0)$ or $\psi_{12}(\mathbf{s}, m_0)$, which should be adjusted in estimation of $\psi_{21}(m_0)$ and $\psi_{22}(m_0)$. As the second term approaches to 0 in probability according to Theorem 3 given below, such an adjustment does not affect the large-sample property of an estimator of $\theta_{m_0}(\mathbf{s})$, which is given by

$$\hat{\theta}_{m_0}(\mathbf{s}) = \frac{\hat{\psi}_{11}(\mathbf{s}, m_0)\hat{\psi}_{22}(m_0)}{\hat{\psi}_{21}(m_0)\hat{\psi}_{12}(\mathbf{s}, m_0)}. \tag{12}$$

The consistency of $\hat{\theta}_{m_0}(\mathbf{s})$ is considered under the condition that $\kappa \to \infty$, where

$$\kappa = \int_{\mathcal{S}} \int_{\mathcal{M}} \lambda(\mathbf{s}, m) dm d\mathbf{s}$$

is the expected number of events in the whole study area. We denote $\xrightarrow{P}$ as convergence in probability as $\kappa \to \infty$. We say $\hat{\theta}_{m_0}(\mathbf{s})$ is a consistent estimator of $\theta_{m_0}(\mathbf{s})$ if $\hat{\theta}_{m_0}(\mathbf{s}) \xrightarrow{P} \theta_{m_0}(\mathbf{s})$ at every $m_0 \in \mathcal{M}$ and $\mathbf{s} \in \mathcal{S}$.

As the consistency of $\hat{\theta}_{m_0}(\mathbf{s})$ is derived under the condition that $\kappa \to \infty$, we need to assume that the intensity functions of N depends on κ. For convenience, we write λ as λ_κ for the first-order intensity function, λ_k as $\lambda_{k,\kappa}$ for the kth-order intensity function, and g as g_κ for pair correlation function. To simplify our proof, we write $\lambda_k(\mathbf{s}, m) = 0$ if $\mathbf{s} \notin \mathcal{S}$. To prove the consistency of $\hat{\theta}_{m_0}(\mathbf{s})$, we need the following regularity conditions:

(C1) The kernel function $K(\mathbf{s})$ is nonnegative in an open neighborhood of origin. It is continuous at any $\mathbf{s}$ if $K(\mathbf{s}) > 0$, and decreasing in any direction as $\mathbf{s}$ moves away from the origin.
(C2) The bandwidth satisfies $h \to 0$ and $\kappa h^d \to \infty$ as $\kappa \to \infty$.
(C3) $\mathcal{S}$ is an open bounded measurable subset of R^d and $\mathcal{M}$ is an open finite sub-interval of R.
(C4) $\lambda_0(\mathbf{s}, m) = \lim_{\kappa\to\infty} \kappa^{-1}\lambda_\kappa(\mathbf{s}, m)$ positively exists and is continuous in $\mathbf{s}$ and m.
(C5) There exists an $H_1(\mathbf{s}, m)$ such that $\kappa^{-1}\lambda_\kappa(\mathbf{s}, m) \leq H_1(\mathbf{s}, m)$ with

$$\int_{\mathcal{S}} \int_{\mathcal{M}} H_1(\mathbf{s}, m) dm d\mathbf{s} < \infty$$

for every $\mathbf{s} \in \mathcal{S}$ and $m \in \mathcal{M}$.
(C6) $\lim_{\kappa\to\infty} g_\kappa[(\mathbf{s}_1, m_1), (\mathbf{s}_2, m_2)] = 1$ if $(\mathbf{s}_1, m_1) \neq (\mathbf{s}_2, m_2)$.
(C7) There is an $H_2[(\mathbf{s}_1, m_1), (\mathbf{s}_2, m_2)]$ such that $|g_\kappa[(\mathbf{s}_1, m_1), (\mathbf{s}_2, m_2)] - 1| \leq H_2[(\mathbf{s}_1, m_1), (\mathbf{s}_2, m_2)]$ with

$$\int_{\mathcal{S}} \int_{\mathcal{M}} \int_{\mathcal{S}} \int_{\mathcal{M}} H_2[(\mathbf{s}_1, m), (\mathbf{s}_2, m)] H_1(\mathbf{s}_1, m_1) H_1(\mathbf{s}_2, m_2) dm_2 d\mathbf{s}_2 dm_1 d\mathbf{s}_1 < \infty$$

for every $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{S}$ and $m_1, m_2 \in \mathcal{M}$.

Lemma 1. *Let* $U_{\mathbf{s}} = \{\mathbf{s}' : \|\mathbf{s} - \mathbf{s}'\| \leq h\}$ *and assume Conditions* (A1)–(A3) *hold. If Conditions* (C2)–(C5) *also hold, then*

$$\lim_{\kappa\to\infty} \theta_{m_0}(U_{\mathbf{s}}) = \lim_{\kappa\to\infty} \theta_{m_0}(\mathbf{s}) = \frac{\int_{\underline{m}}^{m_0} \lambda_0(\mathbf{s}, m) dm \int_{\mathcal{S}} \int_{m_0}^{\overline{m}} \lambda_0(\mathbf{s}, m) dm d\mathbf{s}}{\int_{m_0}^{\overline{m}} \lambda_0(\mathbf{s}, m) dm \int_{\mathcal{S}} \int_{\underline{m}}^{m_0} \lambda_0(\mathbf{s}, m) dm d\mathbf{s}}.$$

Proof. Condition (C2) implies that $|U_{\mathbf{s}}|$ goes to $\mathbf{s}$ as $\kappa \to \infty$. Condition (C5) implies the Dominated Convergence Theorem. Then, we have

$$\begin{aligned}\lim_{\kappa\to\infty}\frac{1}{\kappa|U_{\mathbf{s}}|}E(n_{11}) &= \lim_{\kappa\to\infty}\frac{1}{\kappa|U_{\mathbf{s}}|}\int_{\underline{m}}^{m_0}\int_{U_{\mathbf{s}}}\lambda_\kappa(\mathbf{s}', m)d\mathbf{s}'dm\\ &= \int_{\underline{m}}^{m_0}\lim_{\kappa\to\infty}\frac{1}{|U_{\mathbf{s}}|}\int_{U_{\mathbf{s}}}\frac{\lambda_\kappa(\mathbf{s}', m)}{\kappa}d\mathbf{s}'dm\\ &= \int_{\underline{m}}^{m_0}\lambda_0(\mathbf{s}, m)dm.\end{aligned}$$

Using the same method for n_{12}, n_{21}, and n_{22}, we have

$$\lim_{\kappa\to\infty}\frac{1}{\kappa|U_{\mathbf{s}}|}E(n_{12}) = \int_{m_0}^{\overline{m}}\lambda_0(s, m)dm,$$

$$\lim_{\kappa\to\infty}\frac{1}{\kappa}E(n_{21}) = \int_{\mathcal{S}}\int_{\underline{m}}^{m_0}\lambda_0(s, m)dm$$

and

$$\lim_{\kappa\to\infty}\frac{1}{\kappa}E(n_{22}) = \int_{\mathcal{S}}\int_{m_0}^{\overline{m}}\lambda_0(s, m)dm,$$

which is enough to draw the conclusion. ◇

According to Lemma 1, Conditions (C2)–(C5) are assumptions of the existence of the limits of $\theta_{m_0}(U_{\mathbf{s}})$ and $\theta_{m_0}(\mathbf{s})$. With Conditions (C1), (C6), and (C7), the consistency of $\hat{\theta}_{m_0}(\mathbf{s})$ is derived. In our justification, we find that these conditions are reasonable and weak. They are satisfied at least by a marked Poisson process because it has $g_k[(\mathbf{s}_1, m_1), (\mathbf{s}_2, m_2)] = 1$ for all $\mathbf{s}_1, \mathbf{s}_2 \in \mathcal{S}$ and $m_1, m_2 \in \mathcal{M}$. If N is not a marked Poisson process but it satisfies some additional assumptions (such as the strong mixing condition given by Ivanoff (1982)) as $\kappa \to \infty$, then Conditions (C6) and (C7) also hold.

Theorem 3. *If Conditions* (A1)–(A3) *and* (C1)–(C7) *hold, then* $\hat{\theta}_{m_0}(\mathbf{s}) - \theta_{m_0}(\mathbf{s}) \xrightarrow{P} 0$ *at every* $\mathbf{s} \in \mathcal{S}$ *and* $m_0 \in \mathcal{M}$ *as* $\kappa \to \infty$.

Proof. To show the conclusion, we need to derive the mean and variance formulae of $\hat{\psi}_{11}(\mathbf{s}, m_0)$, $\hat{\psi}_{12}(\mathbf{s}, m_0)$, $\hat{\psi}_{21}(\mathbf{s}, m_0)$, and $\hat{\psi}_{22}(\mathbf{s}, m_0)$. To derive the formulae, we can use the idea for intensity functions of counting processes on real numbers (Ramlau-Hansen, 1983). As all of the derivations are similar, we only display the detail for $\hat{\psi}_{11}(\mathbf{s}, m_0)$ as follows. Simply extending the idea of the proof in Proportion 3.2.1 in Ramlau-Hansen (1983) from R to R^d, we have

$$E\left[\frac{1}{\kappa}\hat{\psi}_{11}(\mathbf{s}, m_0)\right] = \int_{R^d}\int_{\underline{m}}^{m_0}K(\mathbf{s}')\frac{\lambda_\kappa(\mathbf{s} - h\mathbf{s}', m)}{\kappa}dmd\mathbf{s}'$$

and

$$\begin{aligned}V\left[\frac{1}{\kappa}\hat{\psi}_{11}(\mathbf{s}, m_0)\right] &= \int_{R^d}\int_{\underline{m}}^{m_0}\int_{R^d}\int_{\underline{m}}^{m_0}K(\mathbf{s}')K(\mathbf{s}''))\{g_k[(\mathbf{s} - h\mathbf{s}', m'), (\mathbf{s} - h\mathbf{s}'', m'')] - 1\}\\ &\quad\times\frac{\lambda_k(\mathbf{s} - h\mathbf{s}', m')}{\kappa}\frac{\lambda_k(\mathbf{s} - h\mathbf{s}'', m'')}{\kappa}dm''d\mathbf{s}''dm'd\mathbf{s}'\\ &\quad+\frac{1}{\kappa h^d}\int_{R^d}\int_{\underline{m}}^{m_0}K^2(\mathbf{s}')\frac{\lambda_\kappa(\mathbf{s} - h\mathbf{s}', m)}{\kappa}dmd\mathbf{s}'.\end{aligned}$$

Note that Conditions (C5) and (C7) imply the Dominated Convergence Theorem. The first term in the variance expression above goes to 0 as $\kappa \to \infty$ if Condition (C6) hold. Note that Condition (C1) implies

$K(\mathbf{s}')$ is bounded. The second term goes to 0 as $\kappa \to \infty$ if Conditions (C1) and (C5) hold. Note that Condition (C1) also implies the integrand in mean expression above goes to $\lambda_0(\mathbf{s}, m)$ for any fixed $\mathbf{s}' \in \mathcal{S}$. With Conditions (C4) and (C5), we have

$$\frac{1}{\kappa}\hat{\psi}_{11}(\mathbf{s}, m_0) \xrightarrow{P} \int_{R^d}\int_{\underline{m}}^{m_0} K(\mathbf{s}')\lambda_0(\mathbf{s}, m)dmd\mathbf{s}' = \int_{\underline{m}}^{m_0} \lambda_0(\mathbf{s}, m)dm.$$

Similarly, we can show

$$\frac{1}{\kappa}\hat{\psi}_{12}(\mathbf{s}, m_0) \xrightarrow{P} \int_{m_0}^{\overline{m}} \lambda_0(\mathbf{s}, m)dm.$$

These two equations also imply the second term in (10) or (11) goes to zero in probability. Therefore, it is enough to focus on the first term in (10) or (11). Using the same method, we have

$$\frac{1}{\kappa}\hat{\psi}_{21}(m_0) \xrightarrow{P} \int_{R^d}\int_{\underline{m}}^{m_0} \lambda_0(\mathbf{s}, m)dmd\mathbf{s}$$

and

$$\frac{1}{\kappa}\hat{\psi}_{22}(m_0) \xrightarrow{P} \int_{R^d}\int_{m_0}^{\overline{m}} \lambda_0(\mathbf{s}, m)dmd\mathbf{s}.$$

Using the Continuous Mapping Theorem, we have $\hat{\theta}_{m_0}(\mathbf{s}) - \theta_{m_0}(\mathbf{s}) \xrightarrow{P} 0$ at every m_0 and $\mathbf{s}$. ◇

In applications, we may choose $K(\mathbf{s})$ as a Gaussian kernel or uniform kernel. The Gaussian kernel is defined by

$$K_G(\mathbf{s}) = \frac{1}{(2\pi)^{\frac{d}{2}}} e^{-\frac{\|\mathbf{s}\|^2}{2}}$$

and a consistent estimator is derived using Eq. (12) if Conditions (C2)–(C7) hold. A uniform kernel is defined by

$$K_u(\mathbf{s}) = \begin{cases} 1/\pi, & \text{if } \|\mathbf{s}\| < 1, \\ 0, & \text{otherwise.} \end{cases}$$

According to the uniform kernel, there are $\hat{\psi}_{11}(\mathbf{s}, m_0) = n_{11}/h^d = \#\{(\mathbf{s}_i, m_i) : \|\mathbf{s} - \mathbf{s}_i\| < h, m_i < m_0\}/h^d$, $\hat{\psi}_{12}(\mathbf{s}, m_0) = n_{12}/h^d = \#\{(\mathbf{s}_i, m_i) : \|\mathbf{s} - \mathbf{s}_i\| < h, m_i \geq m_0\}/h^d$, $\hat{\psi}_{21}(\mathbf{s}, m_0) = n_{21} = \#\{(\mathbf{s}_i, m_i) : \|\mathbf{s} - \mathbf{s}_i\| \geq h, m_i < m_0\}$, and $\hat{\psi}_{22}(\mathbf{s}, m_0) = n_{22} = \#\{(\mathbf{s}_i, m_i) : \|\mathbf{s} - \mathbf{s}_i\| \geq h, m_i \geq m_0\}$. Then,

$$\hat{\theta}_{m_0}(\mathbf{s}) = \frac{\hat{\psi}_{11}(\mathbf{s}, m_0)\hat{\psi}_{22}(m_0)}{\hat{\psi}_{21}(m_0)\hat{\psi}_{12}(\mathbf{s}, m_0)} = \frac{n_{11}n_{22}}{n_{12}n_{21}}. \tag{13}$$

According to Theorem 3, $\hat{\theta}_{m_0}$ is a consistent estimator if Conditions (C2)–(C7) hold. In the following sections, we numerically study the performance of the kernel estimator of $\theta_{m_0}(\mathbf{s})$ given by (12) based on the Gaussian kernel and the uniform kernel, respectively.

4. Simulation

We simulated realizations from both a marked Poisson process and marked Poisson cluster process on the unit disk on R^2 such that $\mathcal{S} = \{(s_x, s_y) : s_x^2 + s_y^2 \leq 1\}$. We chose these processes because of their popularity in modeling ecological and environmental data. In order to generate a marked Poisson process or a marked Poisson cluster process, we first generate the underlying Poisson point process or Poisson cluster process N_s and then generated the corresponding marks based on conditional distributions at given points.

For both processes, the first-order intensity function of N_s was defined as

$$\lambda_a(\mathbf{s}) = \begin{cases} \kappa a e^{-a\|\mathbf{s}\|^2}/[\pi(1-e^{-a})], & \text{if } a > 0, \\ \kappa/\pi, & \text{if } a = 0, \end{cases} \tag{14}$$

provided $\|\mathbf{s}\| \leq 1$, where κ was pre-selected constant. We chose the above model because it yielded $E[N_s(\mathcal{S})] = \kappa$, which made the point density as

$$f_a(\mathbf{s}) = \lambda_a(\mathbf{s})/\kappa.$$

To successfully generate the underlying N_s, for marked Poisson process, we generated homogeneous Poisson process on $\mathcal{S}$ with the intensity function equal to $\tilde{\lambda}_a(\mathbf{s}) = \lambda_a(\mathbf{s})a/(1-e^{-a})$. As long as points were derived, each point was then thinned with the retained probability equal to $e^{-a\|\mathbf{s}\|}$ which produced a Poisson point N_s with intensity $\lambda_a(\mathbf{s})$. For the marked Poisson cluster process, we first generated homogeneous parent process with intensity function equal to $\tilde{\lambda}_a(\mathbf{s})/k$. As long as the parent process was derived, we generated the offspring process according to the rule of Poisson cluster process (Diggle et al., 2003): the number of offspring points corresponding to each parent point followed a Poisson distribution with mean k and the position of each offspring relative to its parent was defined as a radial symmetric Gaussian random variable with standard deviation σ. We chose $\sigma = 0.02$. After that, each offspring was then thinned according to the same method that we used for the marked Poisson process. Therefore in both cases, we had the first-order intensity function of N_s equal to $\lambda_a(\mathbf{s})$ given by Eq. (14). As long as points were derived, we generated marks independently from $N(\nu, 1)$ distribution with

$$\nu = \nu(\mathbf{s}) = \eta(1-\|\mathbf{s}\|), \quad \eta \geq 0. \tag{15}$$

It was clear that both stationary and nonstationary N_s could be generated respectively according to Eqs. (14) and (15): a stationary N_s was derived if $a = 0$ and a nonstationary N_s was derived if $a \neq 0$. As long as points were derived, we had $\theta_{m_0}(\mathbf{s}) = 1$ for any m_0 and $\mathbf{s}$ if and only if $\eta = 0$.

According to the statistical model in our simulation, we had $\psi_{11}(\mathbf{s}, m_0) = \Phi[m_0 - \eta(1-\|\mathbf{s}\|)]\lambda_a(\mathbf{s})$, $\psi_{12}(\mathbf{s}, m_0) = 1 - \Phi[m_0 - \eta(1-\|\mathbf{s}\|)]\lambda_a(\mathbf{s})$. If $a \neq 0$, then

$$\begin{aligned} \frac{1}{\kappa}\psi_{21}(m_0) &= \frac{1}{1-e^{-a}}[\Phi(m_0-\eta) - e^{-a}\Phi(m_0)] \\ &\quad + \frac{\eta e^{-\frac{a(m_0-\eta)^2}{\eta^2+2a}}}{(1-e^{-a})\sqrt{\eta^2+2a}}\left\{\Phi\left[\sqrt{\eta^2+2a}\left(1+\frac{\eta(m_0-\eta)}{\eta^2+2a}\right)\right] - \Phi\left[\frac{\eta(m_0-\eta)}{\sqrt{\eta^2+2a}}\right]\right\} \end{aligned}$$

and

$$\begin{aligned} \frac{1}{\kappa}\psi_{22}(m_0) &= \frac{1}{1-e^{-a}}[(1-e^{-a}) + \Phi(m_0)e^{-a} - \Phi(m_0-\eta)] \\ &\quad - \frac{\eta e^{-\frac{a(m_0-\eta)^2}{\eta^2+2a}}}{(1-e^{-a})\sqrt{\eta^2+2a}}\left\{\Phi\left[\sqrt{\eta^2+2a}\left(1+\frac{\eta(m_0-\eta)}{\eta^2+2a}\right)\right] - \Phi\left[\frac{\eta(m_0-\eta)}{\sqrt{\eta^2+2a}}\right]\right\}. \end{aligned}$$

If $a = 0$, then

$$\begin{aligned} \frac{1}{\kappa}\psi_{21}(m_0) &= \Phi(m_0) - \frac{1}{\eta^2}\left\{[1+(\eta-m_0)^2][\Phi(m_0) - \Phi(m_0-\eta)]\right. \\ &\quad \left. + \frac{\eta-m_0}{\sqrt{2\pi}}e^{-\frac{(m_0-\eta)^2}{2}} + \frac{m_0-2\eta}{\sqrt{2\pi}}e^{-\frac{m_0^2}{2}}\right\} \end{aligned}$$

and

$$\begin{aligned} \frac{1}{\kappa}\psi_{22}(m_0) &= [1-\Phi(m_0)] + \frac{1}{\eta^2}\left\{[1+(\eta-m_0)^2][\Phi(m_0) - \Phi(m_0-\eta)]\right. \\ &\quad \left. + \frac{\eta-m_0}{\sqrt{2\pi}}e^{-\frac{(m_0-\eta)^2}{2}} + \frac{m_0-2\eta}{\sqrt{2\pi}}e^{-\frac{m_0^2}{2}}\right\}. \end{aligned}$$

Therefore, the value of $\theta_{m_0}(\mathbf{s})$ could be analytically derived.

Table 2

Simulated values of $\widehat{MISE}$ between $\log[\hat{\theta}_{m_0}(\mathbf{s})]$ and $\log[\theta_{m_0}(\mathbf{s})]$ for selected a, η, m_0, and h with respect to the normal kernel K_n and the uniform kernel K_u in the marked Poisson and marked Poisson cluster processes, where the bandwidth was chosen as $h = \omega/\kappa^{1/5}$ with $\omega = 1, 2$.

m_0	a	η	κ	Marked poisson process				Marked poisson cluster process			
				K_n for diff ω		K_u for diff ω		K_n for diff ω		K_u for diff ω	
				1	2	1	2	1	2	1	2
1	0	0	1000	0.0384	0.0084	0.1774	0.0524	0.0376	0.0083	0.1920	0.0537
			5000	0.0136	0.0033	0.0556	0.0160	0.0135	0.0034	0.0585	0.0163
		1	1000	0.0555	0.1198	0.1263	0.0612	0.0578	0.1203	0.1461	0.0643
			5000	0.0224	0.0726	0.0423	0.0219	0.0228	0.0733	0.0454	0.0228
		2	1000	0.1462	0.4414	0.1161	0.1349	0.1587	0.4456	0.1489	0.1473
			5000	0.0628	0.2815	0.0418	0.0557	0.0654	0.2831	0.0468	0.0584
	1	0	1000	0.0373	0.0084	0.1794	0.0522	0.0383	0.0084	0.1938	0.0545
			5000	0.0136	0.0033	0.0560	0.0160	0.0138	0.0035	0.0588	0.0164
		1	1000	0.0791	0.1574	0.1389	0.0938	0.0829	0.1585	0.1571	0.0982
			5000	0.0423	0.1061	0.0556	0.0452	0.0426	0.1067	0.0579	0.0461
		2	1000	0.2280	0.5618	0.1702	0.2450	0.2352	0.5638	0.1989	0.2531
			5000	0.1298	0.3872	0.0912	0.1330	0.1301	0.3884	0.0936	0.1355
2	0	0	1000	0.1448	0.0496	0.3842	0.3464	0.1497	0.0525	0.4312	0.3491
			5000	0.0791	0.0210	0.3312	0.1114	0.0769	0.0220	0.3237	0.1147
		1	1000	0.1452	0.2255	0.3550	0.2028	0.1470	0.2257	0.3686	0.2096
			5000	0.0620	0.1383	0.1992	0.0700	0.0627	0.1389	0.2030	0.0709
		2	1000	0.2713	0.7904	0.2828	0.2876	0.2811	0.7959	0.3222	0.3094
			5000	0.1205	0.5207	0.1365	0.1179	0.1241	0.5236	0.1456	0.1220
	1	0	1000	0.1433	0.0513	0.3808	0.3496	0.1502	0.0498	0.4288	0.3423
			5000	0.0775	0.0218	0.3294	0.1140	0.0791	0.0220	0.3259	0.1146
		1	1000	0.1953	0.3039	0.3878	0.2583	0.2008	0.3046	0.4061	0.2611
			5000	0.0971	0.2088	0.2081	0.1098	0.1001	0.2089	0.2116	0.1112
		2	1000	0.4436	1.0658	0.3740	0.5045	0.4466	1.0695	0.4128	0.5250
			5000	0.2547	0.7665	0.2002	0.2715	0.2583	0.7672	0.2069	0.2748

In addition to the analytical computation of $\theta_{m_0}(\mathbf{s})$, we also numerically computed $\hat{\theta}_{m_0}(\mathbf{s})$ according to the Gaussian kernel and the uniform kernel, respectively. We compared the performance of $\hat{\theta}_{m_0}(\mathbf{s})$ according to the mean integrated square error (MISE) between the logarithm of $\hat{\theta}_{m_0}(\mathbf{s})$ and $\theta_{m_0}(\mathbf{s})$ given by

$$MISE = E\{\log[\hat{\theta}_{m_0}(\mathbf{s})] - \log[\theta_{m_0}(\mathbf{s})]\}^2 = \int_{\|\mathbf{s}\|\le 1} \{\log[\hat{\theta}_{m_0}(\mathbf{s})] - \log[\theta_{m_0}(\mathbf{s})]\}^2 f_a(\mathbf{s})d\mathbf{s}.$$

To estimate the MISE, we evenly selected L points $\{\mathbf{s}_{0i} : i = 1, \ldots, L\}$ in $\mathcal{S}$ and computed

$$\widehat{MISE} = \left\{\sum_{i=1}^{L} f_a(\mathbf{s}_{0i})\right\}^{-1} \sum_{i=1}^{L} \{\log[\hat{\theta}_{m_0}(\mathbf{s}_{0i})] - \log[\theta_{m_0}(\mathbf{s}_{0i})]\}^2 f_a(\mathbf{s}_{0i}).$$

We used $\widehat{MISE}$ to estimate $MISE$ and compared the performance of $\theta_{m_0}(\mathbf{s})$ for selected m_0, a, η, and κ with respect to the Gaussian kernel K_n and the uniform kernel K_u, respectively. According to the asymptotic property of kernel density estimation (van der Vaart, 1998, page 344), the bandwidth h was selected to be proportional to $\kappa^{-1/5}$ such that the expected number of points in a disk with radius h was also approximately proportional to $\kappa^{-1/5}$, which yielded an optimal $MISE$ of order $\kappa^{-4/5}$. Therefore, the bandwidth in our simulation had the form of $h = \omega/\kappa^{1/5}$ for a constant ω.

We simulated 1000 realizations using each set of parameters for the two types of processes. For each realization, we applied the proposed method to estimate $\theta_{m_0}(\mathbf{s})$ and computed $\widehat{MISE}$ using the normal kernel K_n and the uniform kernel K_u, respectively. Specifically, we chose $\mathbf{s}_{0i}$ as points on a 51×51 lattice defined by $\{(x, y) : x = \pm 0.04j_1, y = \pm 0.04j_2, 0 \le j_1, j_2 \le 25\}$. We excluded those $\mathbf{s}_{0i}$ if $\|\mathbf{s}_{0i}\| > 1$, which yielded $L = 1961$ points in the computation of $\widehat{MISE}$.

Table 2 displays the simulated values of $\widehat{MISE}$ between $\log[\hat{\theta}_{m_0}(\mathbf{s})]$ and $\log[\theta_{m_0}(\mathbf{s})]$ in the marked Poisson process and marked Poisson cluster process, respectively. It showed that $\widehat{MISE}$ decreased as

κ increased. This was expected because a larger κ value indicated a more precise estimate of $\theta_{m_0}(\mathbf{s})$. The values of $\widehat{MISE}$ were significantly affected by m_0, η and ω. We compared the performance of $\widehat{MISE}$ between marked Poisson process and marked Poisson cluster process as well as the values of a and found that values of $\widehat{MISE}$ were close. Therefore, we concluded that the performance of $\hat{\theta}_{m_0}(\mathbf{s})$ was significantly affected by the first-order property but not by the second-order property of N_s. In the comparison between kernel functions, we found that the performance of Gaussian kernel was better if η was small (i.e. $\eta = 0$) but worse if η was large (i.e. $\eta = 2$). Although we could not general conclude which kernel function should be better used in applications, we might conclude that the performance of uniform kernel was better if point patterns were inhomogeneous. Overall, we concluded that the influence of m_0, η, κ, bandwidth, and kernel functions was substantial but the dependence between points in N_s was not.

5. Case study

We applied our approach to the *Alberta Forest Wildfire* data and the *Japan Earthquake* data. Using the method proposed by Schoenberg (2004) for testing independence between points and marks, we found that the independence between points and marks was rejected in both datasets. The result was not displayed in this article as it was not the interest in our data analysis. In the *Alberta Forest Wildfire* data, we chose $m_0 = 2$ km^2 as a forest wildfire is called large if the area burned is greater than or equal to 2 km^2. In the *Japan Earthquake* data, we chose $m_0 = 6$ because an earthquake is often called large if the magnitude is greater than or equal to 6. As long as m_0 was decided, we used the Gaussian kernel K_n and the uniform kernel K_u to compute $\hat{\theta}_{m_0}(\mathbf{s})$. As long as $\hat{\theta}_{m_0}(\mathbf{s})$ was derived, we used $\hat{\theta}_{m_0}(\mathbf{s})$ to interpret the relationship between points and marks. According to the definition of $\theta_{m_0}(\mathbf{s})$, it is recommended to pay much attention to subregions with low values of $\hat{\theta}_{m_0}(\mathbf{s})$.

5.1. Alberta forest wildfire data

The *Alberta Forest Wildfire* data consisted of forest wildfire activities in Alberta, Canada, from 1931 to 2012. The Canadian Alberta Forest Service initiated the modern era of wildfire record keeping in 1931. Over the years, this fire information has been recorded, stored and made available in different formats. Beginning in 1996, paper copies of wildfire historical information were no longer retained. The wildfire historical data were entered at the field level on the Fire Information Resource Evaluation System (FIRES), which are available at http://www.srd.alberta.ca/Wildfire.

We collected the historical forest wildfire data from 1996 to 2010 from the website. The dataset contained forest wildfire activities with area burned greater than or equal to 0.01 hectares (Fig. 1). The largest wildfire occurred in 2002 at 111.77 longitude west and 55.47 latitude north and had an area burned of 2388.67 km^2. To apply the approach, we investigated the dependence between the fire occurrences and area burned, in which we treated spatial locations of wildfires as points and area burned as marks.

We estimated the surface of $\theta_{m_0}(\mathbf{s})$ by the kernel method given in Eq. (12). We used the spherical distance and chose $m_0 = 2$ km^2 as the threshold in the method. According to this criterion, we classified a fire as a large fire if its area burned was greater than or equal to 2 km^2. We found that large fires had around 1.6% of total fire occurrences but over 98% of total burned area, which indicated that large fires were more important than small fires in the Alberta Forest Wildfire study.

We computed the spherical distance between all of wildfire locations and found that the largest distance between fires was around 1200 km in the data. We attempted to derive the estimate of $\theta_{m_0}(\mathbf{s})$ at every point in the forest area with the Gaussian kernel and the uniform kernel, respectively. To derive a stable estimate, we investigated multiple choices of h in Eq. (12) and found that $h = 100$ km was almost the best one for both kernel functions. Therefore, we used $h = 100$ km in the computation of $\hat{\theta}_{m_0}(\mathbf{s})$. The values of $\hat{\theta}_{m_0}(\mathbf{s})$ were computed at every point in the study area. Finally, the estimated surface of local odds ratios was derived.

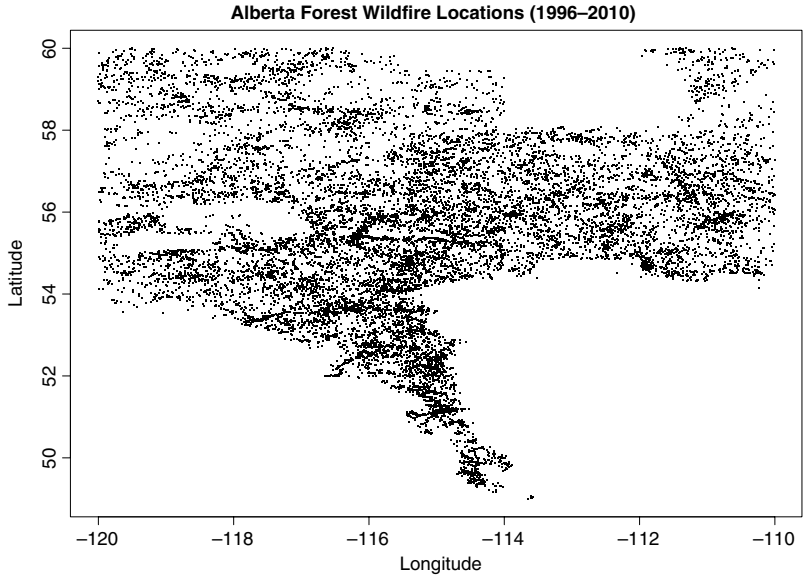

Fig. 1. Wildfire locations in *Alberta forest wildfire* data.

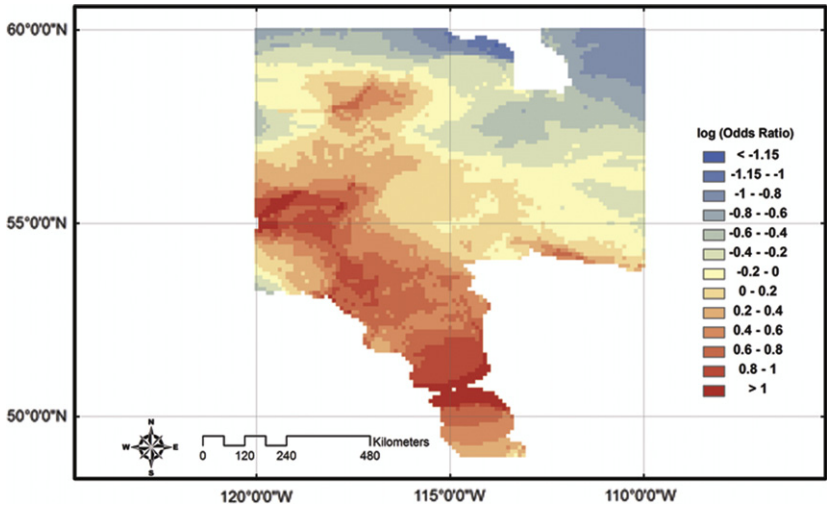

Fig. 2. Estimated surface of $\log[\theta_{m_0}(\mathbf{s})]$ between large ($m \geq m_0 = 2$ km^2) and small ($m < m_0 = 2$ km^2) fires in the *Alberta forest wildfire* data (1996–2010) derived from the uniform kernel with $h = 100$ km.

The estimated surface of local odds ratios from the uniform kernel is displayed in Fig. 2. Although the estimated surface from the Gaussian kernel was similar, we decided only to display the result from the uniform kernel in this paper because point patterns in the *Alberta Forest Wildfire* data seemed inhomogeneous. We found that the values of local odds ratio were mostly small in the North and mostly large in the South, which indicated that the proportion of large fires in the North were higher than the proportion of large fires in the South. Therefore, the risk for a forest wildfire to become a large fire sooner or later in the North was higher than the risk in the South. This is generally consistent with our earlier findings that the forests in the North were with higher risks than in the South using a

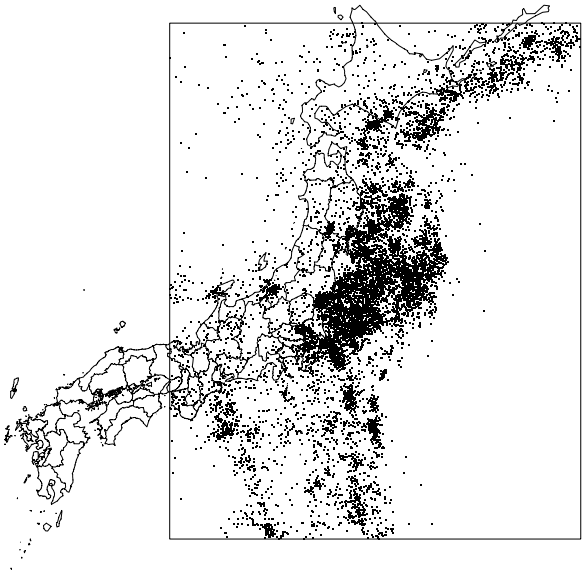

Fig. 3. Occurrences of earthquakes with magnitude greater than or equal to 3.0 in the *Japan Earthquake* data (2002–2011).

power-law analysis approach for fire frequency and area burned of Canadian forest ecosystems (Jiang et al., 2009).

5.2. Japan earthquake data

Many sources of earthquake data are established and available via internet and can be downloaded for free. Examples include the websites of the United States National Geophysical (USGS) data center at http://earthquake.usgs.gov/research/data/, the North California Earthquake Data Center (NCEDC) at http://www.ncedc.org/anss/catalog-search.html, as well as many others. The database contains the date, longitude, latitude, and magnitude of earthquakes at the regional or global level from thousands of years ago to recent years.

We collected the historical earthquake data of the world from the NCEDC website. The collected dataset contained earthquake activities in the world from January 1, 2002 to December 31, 2011 with magnitude greater than or equal to 3.0. We extracted earthquakes in Japan and its nearby ocean area because Japan was considered as the highest risky country for earthquakes in the world. Based on the data we collected, we found that most earthquakes occurred in an area between latitude 30 to latitude 45 north and longitude 130 to 150 east (Fig. 3). Therefore, we decided to take this area as the study region. There were totally 11,423 earthquakes in the dataset. Among those, 130 of them had magnitude between six and seven, 15 of them had magnitude between seven and eight, and 2 of them had magnitude higher than eight. The largest earthquake occurred in the ocean area close to the northeastern Japan in March 11, 2011 at 38.30 latitude north and 142.37 longitude east and had a magnitude of 9.1. The earthquake caused over fifteen thousand human deaths as well as much more injuries. In addition, it also caused a nuclear accident of the Fukushima Daiichi Nuclear Power Plant, which induced a huge leak of nuclear radiation affecting hundreds of thousands of residents and thousands of square kilometer resident area.

We treated the spatial locations of earthquakes as points and magnitude as marks. We estimated the surface of $\theta_{m_0}(\mathbf{s})$ by the kernel method given in (12). We used the spherical distance and chose $m_0 = 6.0$ as the threshold in the method. According to this criterion, we classified an earthquake as large if its magnitude was larger than or equal to 6.0. We derived the estimate of $\theta_{m_0}(\mathbf{s})$ as every point in the study region. To derive a stable estimate, we investigate the choices of h and found that

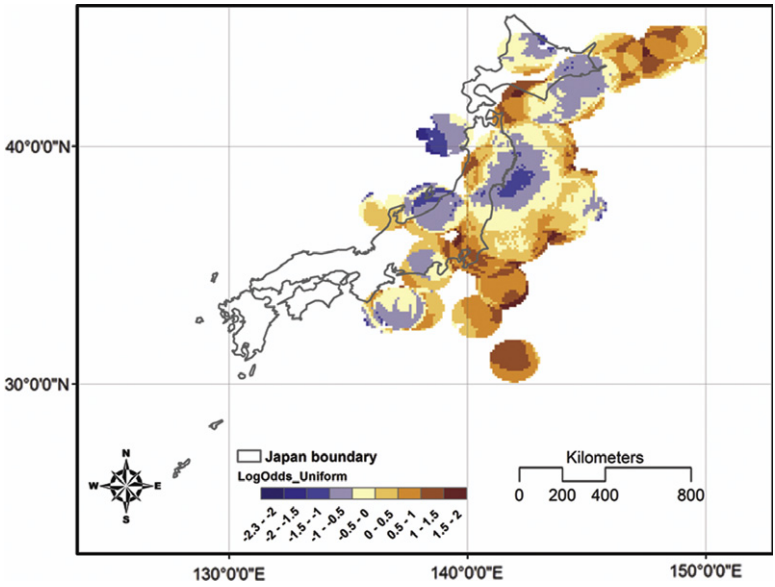

Fig. 4. Estimated surface of $\log[\theta_{m_0}(\mathbf{s})]$ between large ($m \geq m_0 = 6$) and small ($m < m_0 = 6$) earthquakes in the *Japan Earthquake* data (2002–2011) derived from the uniform kernel with $h = 100$ km.

$k = 100$ km was almost the best bandwidth among all the cases that we studied. We used $h = 100$ km in the computation of $\hat{\theta}_{m_0}(\mathbf{s})$ at every $\mathbf{s}$ in the study region. As estimates of $\theta_{m_0}(\mathbf{s})$ at points more than 100 km away from any large earthquake were not stable, we excluded those from our result and finally we got the estimated surface of local odds ratios.

In the comparison between the Gaussian kernel and uniform kernel, we found the main issues were almost the same. Therefore, we decided to only display the results from the uniform kernel (Fig. 4). Another reason for us to focus on the result from the uniform kernel was that point patterns were inhomogeneous in the *Japan Earthquake* data. It showed that values of $\hat{\theta}_{m_0}(\mathbf{s})$ were mostly large in the whole study region except in a few subregions. One subregion was about a circular area centered at around 38.5 latitude north and 142 longitude east. It had high occurrence rates of earthquakes and high proportion of large earthquakes, which was coincident with the largest earthquake, occurred in March 11, 2011. This region was considered as the most risky area for large earthquakes. Another subregion was around 42.5 latitude north and 144 longitude east which also had high occurrence rates and high proportion of large earthquakes. Therefore, it was also considered as a risky area for large earthquakes. A few subregions (e.g. the two in the Sea of Japan) had low occurrence rates of earthquakes but high proportion of large earthquakes, which might still be a concern for large earthquake risks. This was generally consistent with a recent analysis for Japan earthquake hazard maps (Stein et al., 2012). Therefore, we should be able to conclude that the surface of odds ratios displayed in Fig. 4 was able to partially describe hazard patterns for large earthquakes.

6. Discussion

We have proposed an approach of local odds ratio to describe and measure the local dependence between points and marks of MPPs. The approach primarily relies on the definition of the local odds ratio $\theta_{m_0}(\mathbf{s})$ as well as its estimator $\hat{\theta}_{m_0}(\mathbf{s})$. In the definition, there is a threshold value m_0 for marks and a location $\mathbf{s}$ for points. As $\theta_{m_0}(\mathbf{s})$ is considered for every $\mathbf{s}$ in the study area, it is only necessary to determine the value of m_0 in the approach. To decide m_0, we recommend to consider the specific interest in applications such that it can be well explained for events with mark values greater than

and less than m_0, respectively. When m_0 is decided, the estimated value $\hat{\theta}_{m_0}(\mathbf{s})$ at every $\mathbf{s}$ in the study area forms an estimate of the surface of local odds ratios.

According to its definition, the local odds ratio is the ratio of the odds between small and large events at the local level to the global level. It is also equal to the ratio of the relative risk of small events to the relative risks of large events. The value of the local odds ratio describes the risk for an event to finally become large at the local level relative to the global level. The ratio of two local odds ratios describes the risk of large events between the two specific locations. Therefore, the surface of local odds ratios may be used as an index measure for the risk of an event to become a large event in the end conditioning on the observation of an event.

We expect that the proposed approach will find wide applications in natural hazards studies, as many important issues can be discovered by studying the estimated surface of local odds ratios. For example, in a forest wildfire study, if the value of local odds ratio is low, then as long as a fire is observed at the beginning it is more likely for the fire to be large in the end. In an earthquake study, if the value of local odds ratio is low, then energy accumulated is less likely to be released by many small earthquakes rather than by a few large earthquakes. Therefore, our approach suggests that risk studies should pay more attention to the area that has small local odds ratios.

There are a few possible extensions to the approach. First, the local odds ratio defined in this article only involves the first-order intensity function, which cannot be used to interpret higher-order dependence between points and marks. Second, even though we have provided a method to estimate the local odds ratio, we have not considered any explanatory variables. It will be important to develop a method that models the local odds ratio by explanatory variables (e.g., in forest fire studies, the effects of drought condition and fuel loadings). Third, according to the definition, the local odds ratio only involves the odds between events with mark values higher than and lower than a threshold. It cannot be directly used to analyze the dependence between points and marks if two or more critical threshold values are considered. To extend our proposed approach by considering the above issues will be important to wide applications and deserves further investigations.

Acknowledgments

This research is partly supported by the **NASA** Land Use and Land Cover Change program (**NASA**-NNX09AI26G), Department of Energy (**DE**-SC0007007), the United States **NSF**-Division of Information and Intelligent Systems (**NSF**-1028291). The authors appreciate suggestive comments from two anonymous referees which significantly improve the quality of the paper.

References

Agresti, A., 2002. Categorical Data Analysis. Wiley, New York.
Baddeley, A., Turner, R., 2005. spatstat: an R package for analyzing spatial point patterns. J. Statist. Softw. 12, 1–42.
Cressie, N., 1993. Statistics for Spatial Data. Wiley, New York.
Daley, D.J., Vere-Jones, D., 2003. An Introduction to Theory of Point Processes: Volume I: Elementary Theory and Methods, second ed.. Springer, New York.
Diggle, P.J., Ribeiro, J.P.J., Christensen, O., 2003. An introduction to model-based geostatistics. In: Moller, J. (Ed.), Spatial Statistics and Computational Methods, Springer, New York, pp. 43–86.
Guan, Y., Afshartous, D.R., 2007. Test for independence between marks and points of marked point processes: a subsampling approach. Environ. Ecol. Stat. 14, 101–111.
Harte, D., 2010. PtProcess: an R Package for modelling marked point processes indexed by time. J. Statist. Softw. 35, 1–32.
Ho, L.P., Stoyan, D., 2008. Modeling marked point patterns by intensity-marked Cox processes. Statist. Probab. Lett. 78, 2831–2842.
Holden, L., Sannan, S., Bungum, H., 2003. A stochastic marked point process model for earthquakes. Nat. Haz. Earth Syst. Sci. 3, 95–101.
Ivanoff, G., 1982. Central limit theorems for point processes. Stochastic Process. Appl. 12, 171–186.
Jiang, Y., Zhuang, Q., Flannigan, M.D., Little, J.M., 2009. Characterization of wildfire regimes in Canadian boreal terrestrial ecosystems. Internat. J. Wildland Fire 18, 992–1002.
Karr, A.F., 1991. Point Processes and their Statistical Inference, second ed. Marcel Dekker, Inc., New York.
Malamud, B.D., Millington, J.D.A., Perry, G.L.W., 2005. Characterizing wildfire regimes in the United States. In: Proceedings of the National Academy of Sciences of the United States of America, vol. 102, pp. 4694–4699.
Malinowski, A., Schlather, M., Zhang, Z., 2012. Intrinsically weighted means of marked point processes. Arxiv:1201.1335.
McElroy, T., Politis, D.N., 2007. Stable marked point processes. Ann. Statist. 35, 393–419.
Moller, J., Waagepetersen, R., 2007. Modern statistics for spatial point processes. Scand. J. Statist. 34, 643–684.

Myllymäki, M., Penttinen, A., 2009. Conditionally heteroscedastic intensity-dependent marking of log Gaussian Cox processes. Stat. Neerl. 63, 450–473.
Ogata, Y., 1988. Statistical models for earthquake occurrences and residual analysis for point processes. J. Amer. Statist. Assoc. 83, 9–27.
Ogata, Y., Katsura, K., 1993. Analysis of temporal and spatial heterogeneity of magnitude frequency distribution inferred from earthquake catalogs. Geophys. J. Internat. 113, 727–738.
Ogata, Y., 1998. Space–time point process models for earthquake occurrences. Ann. Institut. Statist. Math. 50, 379–402.
Peng, R.D., Schoenberg, F.P, Woods, J.A., 2005. A space–time conditional intensity model for evaluating a wildfire hazard index. J. Amer. Statist. Assoc. 100, 26–35.
Polilitis, D.N., Sherman, M., 2001. Moment estimation for statistics from marked point processes. J. Roy. Statist. Soc. Ser. B 63, 261–275.
Ramlau-Hansen, R., 1983. Smoothing counting process intensities by means of kernel functions. Ann. Statist. 2, 453–466.
Schlather, M., Ribeiro, P.J., Diggle, P.J., 2004. Detecting dependence between marks and locations of marked point processes. J. Roy. Statist. Soc. Ser. B 66, 79–93.
Schoenberg, F.P., 2004. Testing separability in spatial–temporal marked point processes. Biometrics 60, 471–481.
Stein, A., Van Lieshout, M.N.M., Booltink, H.W.G., 2011. Spatial interaction of methylene blue stained soil pores. Geoderma 102, 101–121.
Stein, S., Geller, R.J., Liu, M., 2012. Why earthquake hazard maps often fail and what to do about it. Tectonophysics 562–563, 1–25.
van der Vaart, A.W., 1998. Asymptotic Statistics. Cambridge University Press, Cambridge, UK.
Vere-Jones, D., 1995. Statistical methods for the description and display of earthquake catalogs. In: Walden, A., Guttorp, P. (Eds.), Statistics in Environmental and Earth Sciences. Edward Arnold, London, pp. 220–236.