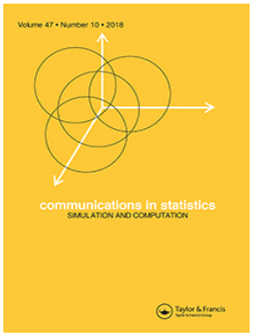

Taylor & Francis
Taylor & Francis Group

Communications in Statistics - Simulation and Computation

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/lssp20

Testing similarity between first-order intensities of spatial point processes. A comparative study

I. Fuentes-Santos, W. González-Manteiga & J. Mateu

To cite this article: I. Fuentes-Santos, W. González-Manteiga & J. Mateu (2021): Testing similarity between first-order intensities of spatial point processes. A comparative study, Communications in Statistics - Simulation and Computation, DOI: 10.1080/03610918.2021.1901118

To link to this article: https://doi.org/10.1080/03610918.2021.1901118

Published online: 19 Mar 2021.

COMMUNICATIONS IN STATISTICS - SIMULATION AND COMPUTATION®
https://doi.org/10.1080/03610918.2021.1901118

Testing similarity between first-order intensities of spatial point processes. A comparative study

I. Fuentes-Santos[a], W. González-Manteiga[b], and J. Mateu[c]

[a]Marine Research Institute, Spanish National Research Council, Vigo, Spain; [b]Department of Statistics, Mathematical Analysis and Optimization, Universidade de Santiago de Compostela, Santiago de Compotela, Spain; [c]Department of Mathematics, University Jaume I, Castellón, Spain

ABSTRACT

Testing whether two spatial point processes have the same spatial distribution is an important task that can be addressed from different perspectives. A Kolmogorov-Smirnov test with asymptotic calibration and a Cramer von Mises type test with bootstrap calibration have recently been developed to compare the first-order intensity of two observed patterns. Motivated by common practice in epidemiological studies, we introduce a regression test based on the relative risk function with two alternative bootstrap calibrations. This paper compares the performance of these nonparametric tests through both an intensive simulation study, and the application to wildfire and crime data. The three tests provide good calibrations of the null hypothesis for simulated Poisson and non-Poisson spatial point processes, but the Cramer von Mises and regression tests outperform the cost-efficient Kolmogorov-Smirnov test in terms of power. In the real data analysis we have seen that the Kolmogorov-Smirnov test does not detect differences between spatial point patterns when dealing with sparse data. In view of these results, it would be preferable using the Cramer von Mises or regression tests despite their higher computational demand.

ARTICLE HISTORY
Received 16 June 2020
Accepted 25 February 2021

KEYWORDS
Bootstrap calibration; Conditional intensity; Criminology; Environmental risk; Inhomogeneous point process; Nonparametric inference

1. Introduction

A common question in the analysis of multitype spatial point processes is whether two types of events have the same spatial structure. This question arises in a wide variety of areas, such as ecology, environmental risk assessment, epidemiology or criminology. We can, for instance, compare the spatial distribution of several species in a given region. In the analysis of environmental risks, we can be interested in assessing if the spatial distribution of arson and natural wildfires in a given region is the same (Fuentes-Santos, González-Manteiga, and Mateu 2017), or whether the spatial pattern of earthquakes in a given region changes after the occurrence of extremely large events (Zhang and Zhuang 2017). Comparison between the spatial distribution of disease cases and the population at risk in order to identify areas with high disease risk has been a major issue for epidemiologists (Kelsall and Diggle 1995; Davies and Hazelton 2010). Differences between the observed patterns imply that they have been generated by different point processes and, consequently, different models should be used to characterize their distribution. In spite of its relevance when dealing with real data, this issue has received little attention, while the analysis of multitype point processes has mainly focused on testing for interactions between different types of events

CONTACT I. Fuentes-Santos isafusa@gmail.com Marine Research Institute, Spanish National Research Council, Vigo, Spain.

through Monte Carlo tests based on second-order characteristics such as the K-cross and the mark correlation functions (Ripley 1981; Diggle 2013). These techniques have been widely used to search for relationships between different types of wildfires (Hering, Bell, and Genton 2009; Juan, Mateu, and Saez 2012; Fuentes-Santos, Marey-Pérez, and González-Manteiga 2013), or to analyze species distribution in plant ecology (Perry, Miller, and Enright 2006; Getzin et al. 2008; Fibich et al. 2016).

A spatial point process is a stochastic process governing the location of a random number of events, $\mathbf{X} = \{\mathbf{x_1}, ..., \mathbf{x_N}\}$, irregularly placed in a planar region $W \subset \mathbb{R}^2$. If each event has associated any measure or mark, we have a marked point process. A multitype point process is a marked point process with categorical marks that define different groups or types of events (Diggle 2013). Throughout this paper, point processes and patterns are denoted in bold capitals, and events are denoted in bold. The spatial distribution of events in Poisson point processes, i.e., those with independent events, is determined by the *first-order intensity function* (Illian et al. 2008; Diggle 2013), which measures the expected number of events per unit area, and is defined as follows

$$\lambda(x) = \lim_{|dx| \to 0} \left\{ \frac{E[N(dx)]}{|dx|} \right\} \tag{1}$$

where $|dx|$ and $N(dx)$ denote the area and number of events of $\mathbf{X}$ in the infinitesimal disk dx centered at location $x \in W$. Events in a spatial point process may not occur independently, and interactions between them are characterized through second-order properties such as the K-function $K(r)$ (Ripley 1977) and the *second-order intensity function*, defined as $\lambda_2(x, y) = \lim_{|dx|, |dy| \to 0} \left\{ E\left[N(dx)N(dy)\right] / |dx||dy| \right\}$. The *conditional intensity function*

$$\lambda_c(x) = \lambda_c(x|\mathbf{y}) = \lambda_2(x, \mathbf{y}) / \lambda(\mathbf{y}) \tag{2}$$

determines the intensity at a point x conditional on the information that there is an event in $\mathbf{y}$ (Diggle 2013), and characterizes uniquely the distribution of events in any spatial point process. In the particular case of a spatial Poisson point process $\lambda_c(x|y) = \lambda(x)$.

Two spatial point patterns with the same spatial structure can be seen as the type i and type j patterns of a random labeled bivariate point process, or as independent realizations of a point process; in both cases their K-functions are equal, $K_i(r) = K_j(r)$ (Ripley 1977; Diggle 2013). This property motivated the use of Monte Carlo tests based on the K-function to check the similarity between two observed patterns (Hahn 2012). However, differences between the K-functions of two inhomogeneous spatial point patterns can reflect differences in the first-order intensities or in the dependence structure, and consequently these tests can lead to wrong conclusions.

This knowledge gap has been recently addressed from different perspectives. Alba-Fernández et al. (2016) and Andresen (2009) introduced area-based tests, which count the number of events within predefined spatial units (e.g., quadrants, space-filling curves) to measure local and global discrepancies between two observed point patterns. Zhang and Zhuang (2017) proposed a Kolmogorov-Smirnov test using the absolute difference between the point densities of the observed patterns over a π-system as a discrepancy measure. Fuentes-Santos, González-Manteiga, and Mateu (2017) proposed a Cramer von Mises type statistic that measures the discrepancy between the densities of event locations of the observed patterns, taking into account that the intensities of two spatial Poisson point processes with the same spatial structure are proportional.

A main issue in epidemiological studies is testing for spatial variation in the risk of a given disease, which involves the comparison between the spatial distribution of disease cases and the population at risk, referred as controls. In this line, Kelsall and Diggle (1995) proposed a kernel estimator of the relative risk function for independent and Poisson case and control patterns, defined as the ratio between the first-order intensities of disease cases and the control population (Bithell 1990). They also proposed a Monte Carlo algorithm to generate tolerance contour surfaces under the null hypothesis of equal spatial distribution of disease cases and controls, and a global Monte Carlo homogeneity test. Davies and Hazelton (2010) introduced an adaptive kernel

estimator of the relative risk function, which is asymptotically normal allowing the generation of tolerance contours without the need of Monte Carlo simulations. Fuentes-Santos, González-Manteiga, and Mateu (2018) developed a regression test based on the log-ratio between the intensity of a spatio-temporal point process and the intensity of its spatial component to test for separability. In the same line, we propose a regression test based on the relative risk function to test if two spatial point processes have the same spatial distribution.

Motivated by the fact that there are not currently comparative analysis among these procedures, this work aims to compare the performance of the three tests outlined above in the analysis of Poisson and non-Poisson point processes. The plan of this work is the following. Sec. 2 introduces the Kolmogorov-Smirnov and Cramer von Mises tests, proposed in the literature, and we additionally consider a relative risk based procedure to compare the spatial distribution of inhomogeneous spatial point processes. The performance of these tests is tested through a simulation study in Sec. 3, and through the analysis of two important real data problems, such as wildfire patterns in Galicia (NW Spain) and gunfire violence in Rio de Janeiro (Brazil) in Sec. 4. The paper ends with a brief discussion in Sec. 5 and conclusions in Sec. 6.

2. Nonparametric tests for first-order comparison

Let $\mathbf{X} = \{\mathbf{x_i}\}_{i=1}^{N}$ be a realization of a bivariate inhomogeneous spatial point process observed in a bounded region $W \subset \mathbb{R}^2$, and let $\mathbf{X_1} = \{\mathbf{x_i}\}_{i=1}^{N_1} = \{\mathbf{x_{1,i}}\}_{i=1}^{N_1}$, $\mathbf{X_2} = \{\mathbf{x_{N_1+j}}\}_{j=1}^{N_2} = \{\mathbf{x_{2,j}}\}_{j=1}^{N_2}$, where $N = N_1 + N_2$, be the spatial patterns of type 1 and type 2 events in $\mathbf{X}$. We denote by $\lambda_1(x)$ and $\lambda_2(x)$ the first-order intensity functions of $\mathbf{X}_1$ and $\mathbf{X}_2$, and by $\lambda_{01}(x) = \lambda_1(x)/m_1$, $\lambda_{02}(x) = \lambda_2(x)/m_2$ their densities of event locations (Cucala 2006), where $m_j = \int_W \lambda_j(x), j = 1, 2$ is the mean intensity or expected number of events of each point process.

If $\mathbf{X}_1$ and $\mathbf{X}_2$ have the same spatial distribution, their conditional intensity functions are proportional. Therefore, we can compare the distribution of two observed patterns testing the null hypothesis

$$\mathcal{H}_0 : \lambda_{c,1}(x) = \omega\lambda_{c,2}(x); \ \forall x \in W \tag{3}$$

for some $\omega > 0$. In case of Poisson point processes, whose conditional and first-order intensities are equal, the null hypothesis reduces to

$$\mathcal{H}_0 : \lambda_1(x) = \omega\lambda_2(x); \ \forall x \in W \tag{4}$$

This hypothesis can be tested using nonparametric approaches, such as the Kolmogorov-Smirnov test by Zhang and Zhuang (2017), the Cramer von Mises test by Fuentes-Santos, González-Manteiga, and Mateu (2017), or the regression test based on the relative risk function proposed in this work. The Kolmogorov-Smirnov test compares the spatial structure of two independent, but not necessarily, Poisson point processes, $\mathbf{X_1}$ and $\mathbf{X_2}$. The test statistic maximizes the absolute difference between the empirical point densities over a π-system, avoiding the need of estimating the first-order intensity of each observed pattern. The Cramer von Mises and regression tests are based on kernel estimators of the density of event locations and of the relative risk function of inhomogeneous Poisson point processes. The Poisson assumption is required to guarantee the consistency of the kernel estimators, and to obtain the asymptotic null distribution of the test statistics.

As argued by Zhang and Zhuang (2017), we can ignore the dependence between point patterns, but not the dependence structure of each spatial point process as both regular and clustered patterns are common in real data analysis. Thus, the Poisson assumption in the kernel-based tests may seem quite restrictive. It should be noted that these tests can also be applied to compare the spatial distribution of non-Poisson point processes, in which case these procedures measure the discrepancy between the conditional intensities, i.e., our null hypothesis is (3). Indeed, as we

cannot distinguish between heterogeneity and interaction between events in an observed pattern without additional information (Diggle 2013), common practice is to assume that the spatial point process is Poisson, estimate its intensity function, and then estimate the second-order characteristics to test for independence between events. In this case, acceptation of the null hypothesis implies that the observed patterns have the same spatial distribution but not that they have been generated by the same point process.

2.1. Kolmogorov-Smirnov test

Zhang and Zhuang (2017) developed a nonparametric procedure to test whether two independent but not necessarily Poisson point processes, $\mathbf{X_1}$ and $\mathbf{X_2}$, have the same spatial distribution. The test is based on the fact that the null hypothesis, $\mathcal{H}_0$, implies that there exists an $\omega > 0$ such that $\mathbb{E}[N_1(A)] = \omega\mathbb{E}[N_2(A)]$ for any Borel set $A \in \mathcal{B}(W)$. Considering this property, the authors define the discrepancy measure

$$D_\omega(A) = N_1(A) - \omega N_2(A) \tag{5}$$

Then, under $\mathcal{H}_0$, $\mathbb{E}\big[D_{\omega_0}(A]\big) = 0$ for any Borel set $A \in \mathcal{B}(W)$ and, consequently $|D_{\omega_0}(A)|$ is close to 0. Zhang and Zhuang (2017) proved that a sufficient condition for $\mathcal{H}_0$ to hold is that $\mathbb{E}[D_\omega(A)] = 0$ for some $\omega > 0$ and any $A \in \mathcal{P}(W)$, where $\mathcal{P}$ is a $\pi-$system that generates $\mathcal{B}(W)$. Replacing ω by its empirical estimator, $\hat{\omega} = N_1(W)/N_2(W)$, in expression (5) we obtain the following test statistic

$$\begin{aligned}\hat{D} &= \frac{1}{\zeta}\sup_{A\in\mathcal{P}} D_{\hat{\omega}}(A) \\ &= \frac{1}{\zeta}\sqrt{\frac{N_1(W)N_2(W)}{N_1(W)+N_2(W)}}\sup_{A\in\mathcal{P}}\left|\frac{N_1(A)}{N_1(W)} - \frac{N_2(A)}{N_2(W)}\right|\end{aligned} \tag{6}$$

where ζ is a normalizing constant needed to guarantee the convergence of the null distribution of $\hat{D}$ to a Brownian bridge, and is estimated by

$$\hat{\zeta}^2 = \frac{1}{K-1}\sum_{i=1}^{K}\left[\frac{\big(N_1(W_i) - \hat{N}_1(W_i)\big)^2}{\hat{N}_1(W_i)} + \frac{\big(N_2(W_i) - \hat{N}_2(W_i)\big)^2}{\hat{N}_2(W_2)}\right] \tag{7}$$

where $\hat{N}_1(W_i) = \hat{\omega}(N_1(W_i) - N_2(W_i))/(1+\hat{\omega})$, and $\hat{N}_2(W_i) = \hat{N}_1(W_i)/(1+\hat{\omega})$ for a partition $\{W_i\}_{i=1}^K$ of the observation domain, W. The distribution of the Brownian bridge, $\mathbb{W}_{F,W}(\cdot)$, depends on K and on F, the function used to define the $\pi-$system.

Zhang and Zhuang (2017) introduced two $\pi-$systems that may be useful in practice. For $W = [0, w_1] \times [0, w_2] \subset \mathbb{R}^2$, we can choose $t = F(x) = (x_1/w_1, x_2/w_2)$ for any $x \in W$, which gives the $\pi-$ system $A_t = [0, x_1/w_1] \times [0, x_2/w_2]$. In this case $\mathbb{W}_{F,W}(t)$ is a bidimensional standard pinned Brownian sheet, and the corresponding critical value at significance level $\alpha = 0.05$ is $D_{0,05} = 1.6522$. In the second case W is an arbitrary region, and we choose $t = F(x)$ such that $A_t = [0, 1]$ and $F^{-1}(A_t) = \{x' \in W; F(x') \leq F(x)\}$. Here $\hat{D}$ converges to the standard Brownian motion whose distribution can be estimated through a Taylor expansion. It should be noted that we reject $\mathcal{H}_0$ if D is larger than the critical value, but this procedure does not allow us to accept $\mathcal{H}_0$ otherwise.

2.2. Cramer von Mises test

Let $\mathbf{X}$ be a realization of a bivariate inhomogeneous spatial Poisson point process observed in a bounded region $W \subset \mathbb{R}^2$, and $\mathbf{X_1}$, $\mathbf{X_2}$ are the spatial point patterns of type 1 and type 2 events in $\mathbf{X}$. Conditional to the number of events, $N_j = n_j$, the observed patterns can be seen as random

samples of the bivariate random distributions with densities $\lambda_{0j}(x), j = 1, 2$. Considering this property, Fuentes-Santos, González-Manteiga, and Mateu (2017) extended the nonparametric test developed by Duong, Goud, and Schauer (2012) for multivariate data to the spatial point process framework, and proposed a test statistic based on an L^2-distance measure

$$\hat{T} = \int_W \left(\hat{\lambda}_{01}(x) - \hat{\lambda}_{02}(x)\right)^2 dx = \hat{\psi}_1 + \hat{\psi}_2 - \left(\hat{\psi}_{12} + \hat{\psi}_{21}\right) \tag{8}$$

where $\hat{\psi}_{ij}$ and $\hat{\psi}_i$ are estimators of $\psi_{ij} = \int_W \lambda_{0i}(x)\lambda_{0j}(x)dx$ for $i, j = 1, 2$ and $\psi_i = \int_W \lambda_{0i}(x)^2 dx$, obtained by kernel smoothing with plug-in bandwidth (Chacón and Duong 2010). The null distribution of $\hat{T}$ is asymptotically normal under regularity conditions analogous to those assumed in the classical multivariate distribution framework. However, in view of the slow convergence rate to the normal distribution, Fuentes-Santos, González-Manteiga, and Mateu (2017) developed a smooth bootstrap algorithm to calibrate the test.

Finally, notice that this test can also be used to compare the spatial distribution of non-Poisson point processes, in which case we test the null hypothesis (3), i.e., we compare the conditional instead of the first-order intensity functions.

2.3. Nonparametric test based on the relative risk function

Let $\mathbf{X}$ be a realization of a bivariate inhomogeneous spatial Poisson point process in a bounded region $W \subset \mathbb{R}^2$, and $\mathbf{X_1}$, $\mathbf{X_2}$, the spatial patterns of type 1 and type 2 events in $\mathbf{X}$. The null hypothesis (4) can be rewriten as $\mathcal{H}_0 : \lambda_1(x)/\lambda_2(x) = \omega;\ \forall x \in W$ for some $\omega > 0$, i.e., the relative risk function (Bithell 1990), $r(x) = \lambda_1(x)/\lambda_2(x)$, is spatially invariant. Considering this property, we shall follow the same strategy as in Fuentes-Santos, González-Manteiga, and Mateu (2018) and use a no-effect test that checks the dependence of the log-relative risk function $\rho(x) = \log r(x)$ on event locations to compare the spatial distribution of $\mathbf{X}_1$ and $\mathbf{X}_2$.

To implement this test, we first need to estimate the log-relative risk function. In this case, we estimate $\rho(\cdot)$ as the log-ratio of the kernel densities of event locations, as proposed by Kelsall and Diggle (1995)

$$\hat{\rho}(x) = log\left(\frac{\hat{\lambda}_{01,h_1}(x)}{\hat{\lambda}_{02,h_2}(x)}\right) \tag{9}$$

where the kernel estimators of the marginal densities of event locations, $\hat{\lambda}_{0j,h_j}(x); j = 1, 2$, are given by

$$\hat{\lambda}_{0j,h_j}(x) = \left(p_{h_j}(x)N\right)^{-1} h^{-2} \sum_{i=1}^{N} k\left(h^{-2}(x - \mathbf{x_i})\right) I(N \neq 0) \tag{10}$$

where $p_{h_j}(x)$ is the edge-correction term, and h_j, $j = 1, 2$ the bandwidth parameter for the case and control kernel estimators. We can use different bandwidths in the numerator and denominator of $\hat{\rho}(\cdot)$, but Kelsall and Diggle (1995) showed that the asymptotically optimal estimator with respect to the mean square error is achieved using the same bandwidth for cases and controls. This choice leads to a bias cancelation in regions where $\lambda_{01} = \lambda_{02}$ and simplifies data-driven bandwidth selection. Kelsall and Diggle (1995) also proposed a least-squares cross-validation bandwidth selector that shall be used in this work.

Once estimated the log-relative risk function, we have a regression problem where the log-relative risk function, $Y = \{y_i = \hat{\rho}(\mathbf{x_i}), i = 1, ..., n\}$, is a response variable that may depend on the spatial covariate $X = \{x_i = (x_{i1}, x_{i2}), i = 1, ..., n\}$ comprising the event locations, and we test for the effect of X on Y discriminating between two competing models (Bowman and Azzalini 1997)

$$\mathcal{H}_0 : E\left[y_i|x_i\right] = \mu$$
$$\mathcal{H}_1 : E\left[y_i|x_i\right] = m(x_i)$$

The null model μ can be estimated by the empirical mean $\hat{y} = \sum_{i=1}^{n} y_i$. The alternative model, $m(x)$ is an unknown smooth function, which can be estimated by kernel regression (Nadaraya 1964; Watson 1964)

$$\hat{m}(x) = \hat{m}(x_1, x_2) = \frac{\sum_{i=1}^{n} w_{g_1}(x_{i1} - x_1) w_{g_2}(x_{i2} - x_2) y_i}{\sum_{i=1}^{n} w_{g_1}(x_{i1} - x_1) w_{g_2}(x_{i2} - x_2)} \tag{11}$$

where the kernel, $w(\cdot)$, is a univariate symmetric density function, and $g = (g_1, g_2)$ is the vector of smoothing parameters. Three alternative procedures have been commonly used to select this parameter: (i) bandwidth selector associated to the approximate degrees of freedom, *df*, of the regression errors, (ii) least-squares cross-validation, and (iii) an AICC-based method.

Discrepancy between the null and alternative models is measured through the following generalized test statistic

$$F = \frac{(RSS_0 - RSS_1)/(df_1 - df_0)}{RSS_1/df_1} \tag{12}$$

where RSS_0 and RSS_1 are the residual sum of squares for nonparametric estimators of the null and alternative models, and df_0, df_1 denote their respective degrees of freedom.

In the linear model framework, the residual sums of squares and the test statistic follow, respectively, a χ^2 and an F distribution. However, these properties are not fulfilled in the context of nonparametric regression. Bowman and Azzalini (1997) proposed two procedures to estimate the distribution of F under the null hypothesis: (i) if the errors in the regression model have normal distribution, the null distribution of F is approximated by a shifted and scaled χ^2; (ii) a computationally intensive procedure based on permutation tests is applied otherwise.

The permutation test relies on the fact that under $\mathcal{H}_0$ the pairing of any particular x and y is completely random. This calibration procedure has shown a good performance for the separablity test developed by Fuentes-Santos, González-Manteiga, and Mateu (2018). Here we propose a modified version of this permutation test taking into account that our null hypothesis implies that the data can be seen as a single realization of the unmarked point process $\mathbf{X}$, followed by independent random allocation of events to $\mathbf{X_1}$ and $\mathbf{X_2}$ with probability proportional to their respective mean intensity. Therefore, following the Monte Carlo procedure introduced by Kelsall and Diggle (1995), we propose a calibration that estimates the null distribution of $\hat{F}$ using a random labeling algorithm that computes the test statistic for B realizations of the null hypothesis generated by random reallocation of case/control marks. The empirical p-value of the test is the proportion of simulated F-statistics larger than that obtained from the observed data.

The test can also be calibrated through a smooth bootstrap algorithm analogous to that used for the Cramer von Misss test. The bootstrap calibration is implemented as follows:

(1) Estimate the log-relative risk function and compute the test statistic $\hat{F}_1$ for the observed patterns, $\mathbf{X_1}$ and $\mathbf{X_2}$.
(2) Compute the kernel intensity estimator, $\hat{\lambda}_H(x)$, of the unmarked point process comprising the location of type 1 and type 2 events, $\mathbf{X^0} = \mathbf{X_1} \cup \mathbf{X_2}$.
(3) For $b = 2, \ldots, B$:
    (1) Generate a bivariate spatial point process $\mathbf{X_b} = \{\mathbf{X_{1,b}}, \mathbf{X_{2,b}}\}$ where for $j = 1, 2$, $\mathbf{X_{j,b}}$ are realizations of spatial Poisson point processes with first-order intensity proportional to that of the unmarked pattern and the same number of events as $\mathbf{X}_j$.
    (2) Obtain the bootstrap log-relative risk functions $\hat{\rho}_b^*$ and compute the corresponding test statistic $\hat{F}_b$.

(4) The probability of rejecting the null hypothesis is the proportion of bootstrap statistics $\{\hat{F}_b\}_{b=2}^{B}$ larger than $\hat{F}_1$.

3. Simulation study

We have conducted a simulation study to analyze the performance of the relative risk based test introduced in Sec. 2.3. We first compare the calibration of the null distribution provided by the random labeling and smooth bootstrap algorithms, and analyze the power of the test. In a second stage we compare the regression test with the Kolmogorov-Smirnov and Cramer von Mises tests. These analysis have been conducted through application of the three nonparametric tests to Poisson, clustered and regular inhomogeneous bivariate spatial point processes. The spatial point processes simulated in this work are based on the simulation studies conducted by Zhang and Zhuang (2017) and Fuentes-Santos, González-Manteiga, and Mateu (2017) to check the performance of their respective tests.

Following Zhang and Zhuang (2017) we simulate Poisson, clustered and regular inhomogeneous spatial point processes on the square region $W = [-10, 10]^2$. We generate inhomogeneous Poisson point processes with first-order intensity $\lambda_{\alpha_j}(x) = \lambda_{\alpha_j}(x_1, x_2) = \kappa_j f_{\alpha_j}(x_1) f_{\alpha_j}(x_2)$, where

$$f_{\alpha_j}(t) = \frac{\Gamma(2\alpha)}{\Gamma^2(\alpha)(2v)^{2\alpha-1}}(v^2 - t^2)^{(\alpha-1)}; \ \ |t| < v \tag{13}$$

is the density of the $\beta(\alpha_j, \alpha_j)$ distribution on $[-v, v]$, in this case $v = 10$. κ_j determines the expected number of events of the simulated patterns, in this case $\kappa_j = 500 w_j, j = 1, 2$, with $w = (1/2, 1/2)$, $w = (2/3, 1/3)$ and $w = (9/10, 1/10)$ for balanced and unbalanced designs. Balanced and unbalanced designs were considered to check whether the asymmetry in the size of the point patterns affects the performance of the tests. We have used $\alpha_1 = 2$, and $\alpha_2 = 2, 3, 4$ to generate the null hypothesis and two alternatives with different degrees of deviation from $\mathcal{H}_0$.

Clustered patterns were simulated using a Thomas cluster point process model. We generate an inhomogeneous spatial point pattern of parent points with intensity function $\lambda_{\alpha_j}(x)/\gamma$, and $Poisson(\gamma)$ offsprings randomly placed around each parent event according to a symmetric Gaussian distribution with standard deviation σ, in this case $\gamma = 5$ and $\sigma = 0.5$.

Regular patterns were generated through a simple inhibition point process model. We first generate inhomogeneous Poisson point processes with first-order intensities proportional to $\lambda_{\alpha_j}(x), j = 1, 2$ and we then apply an inhibition radius of $r = 0.2$. Figure 1 shows the first-order intensity of the inhomogeneous Poisson point processes and the conditional intensities of the Thomas cluster and simple inhibition point processes for the null and alternative hypothesis.

Following Fuentes-Santos, González-Manteiga, and Mateu (2017) we generated inhomogeneous spatial Poisson point processes with first-order intensity $\lambda_j(x) = a_j \exp(-3x_2), j = 1, 2$, for any $x = (x_1, x_2) \in \mathbb{R}^2$ on the unit square. Different values of $a_j, j = 1, 2$ were considered to obtain different proportions of type 1 and type 2 events. The alternative hypothesis was generated adding a random number of events uniformly distributed on a subregion of the type 2 point pattern, as follows

$$\begin{cases} \lambda_1(x) = a_1 \exp(-3x_2) \\ \lambda_2(x) = a_2 \left(\exp(-3x_2) + \dfrac{100}{a_2} \epsilon \left(\dfrac{1}{4}\right)^{-2} I_{(x_1, x_2) \in [0.5, 0.75]^2} \right) \end{cases} \tag{14}$$

As above a_1 and a_2 were fixed to generate realizations of a bivariate point process with $m_1 = m_2 = 250$ in the balanced design, $m_1 = 1000/3, m_2 = 500/3$ and $m_1 = 450, m_2 = 50$ in the unbalanced designs. $(1/4)^2$ in expression (14) is the area of the subregion of the unit square where 100ϵ events are added to generate realizations of the alternative hypothesis in the balanced design,

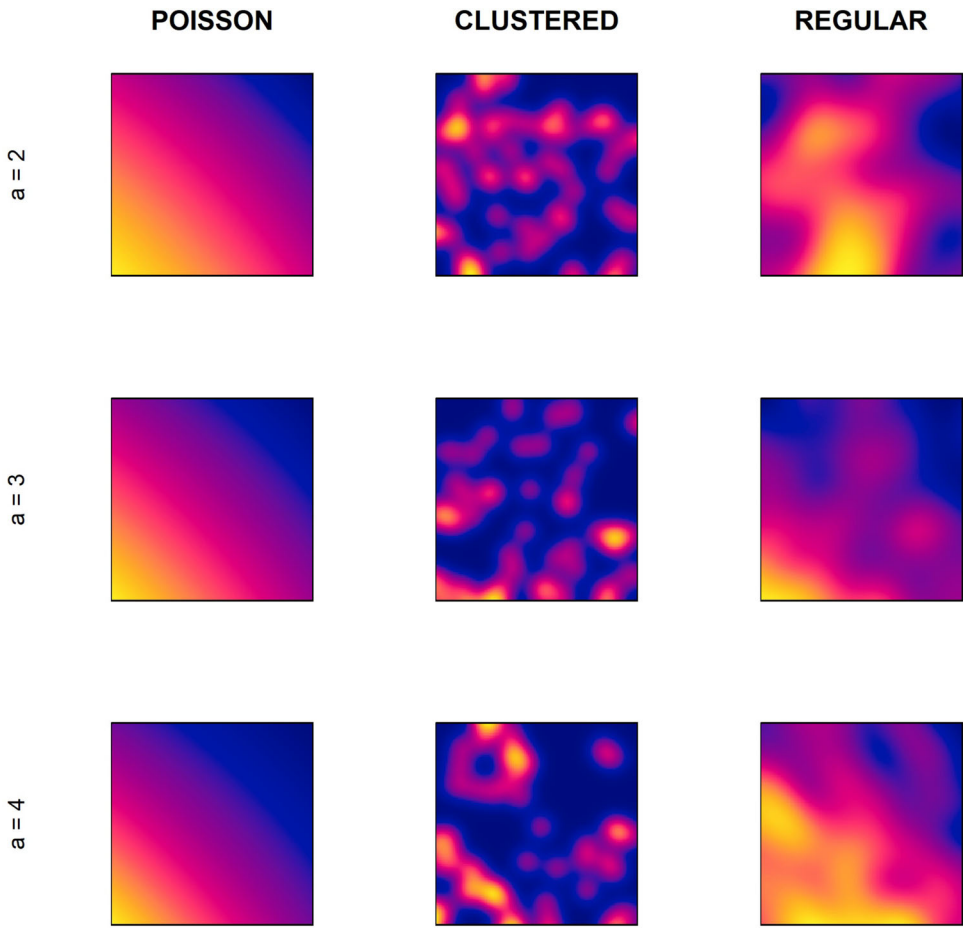

Figure 1. MODEL I: First-order intensities (Poisson) and conditional intensities (clustered and regular) of the type 1 and type 2 events in the bivariate point process based on the intensity function (13). Null hypothesis in the first row.

and $\epsilon = 0.4$, $\epsilon = 0.8$ determines the number of events added to the type 2 point process, that is the degree of departure from the null hypothesis. Clustered patterns were generated as Thomas cluster point processes with $\gamma = 5$, $\sigma = 0.02$ and intensity of the parent points equal to $\lambda_j(x)/\gamma, j = 1, 2$. Regular patterns were generated through simple inhibition point processes with an inhibition radius of $r = 0.02$. Figure 2 shows the first-order intensity of the inhomogeneous Poisson point processes and the conditional intensities of the Thomas cluster and simple inhibition point processes for the null and alternative hypotheses.

For the different scenarios defined throughout this section, we computed the test statistics and the corresponding empirical p-values for 1000 realizations of each bivariate spatial point process. The probability of rejecting the null hypothesis at a given significance level, α, was obtained as the proportion of p-values smaller than α. This simulation study was conducted with the help of the *spatstat* (Baddeley and Turner 2005), *ks* (Duong 2013), *sparr* (Davies, Marshall, and Hazelton 2018) and *sm* (Bowman and Azzalini 2014) packages of the R statistical software (R Core Team 2019).

3.1. Comparison of calibration methods in the regression test

We have applied the regression test (Sec. 2.3) with $B = 200$ realizations of the random labeling and smooth bootstrap calibration procedures to compare the spatial distribution of the different

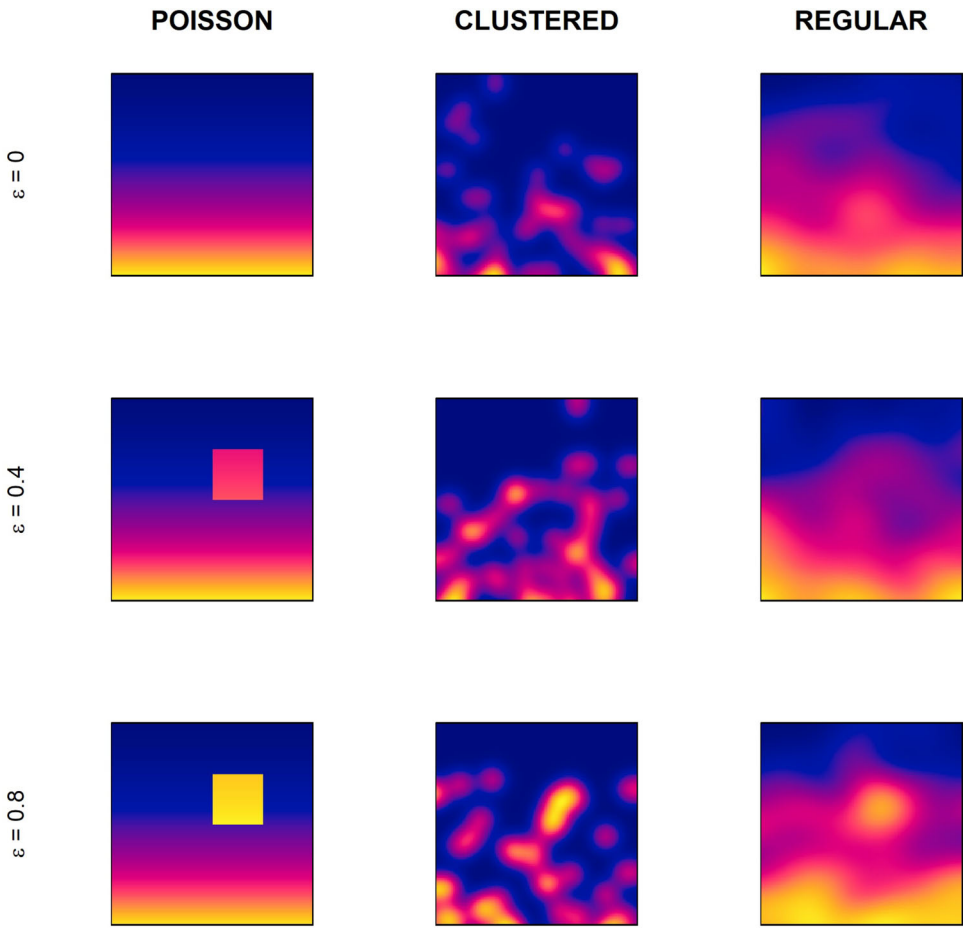

Figure 2. MODEL I1: First-order intensities (Poisson) and conditional intensities (clustered and regular) of the type 1 and type 2 events in the bivariate point process based on the intensity function (14). Null hypothesis in the first row.

bivariate point processes outlined above. The test has been implemented using the least squares cross-validation bandwidth selector in the kernel log-relative risk function (Kelsall and Diggle 1995), and the three bandwidth selectors outlined in Sec. 2.3 in the kernel regression function, $\hat{m}(x)$. Note that in case of asymmetric designs we consider as cases in the relative risk function the point process with lower number of events, $\mathbf{X_2}$.

Table 1 shows the estimated significance levels for Model I under the null hypothesis. Both calibration procedures provide reasonable estimators of the nominal significance levels, although they report different performances for symmetric and asymmetric designs. Random labeling overestimates the size of the test, whereas the bootstrap calibration underestimates it for Poisson point processes. For the Thomas cluster point processes, which have highly inhomogeneous conditional intensities (see Figure 1), random labeling is conservative for the symmetric design but provides good calibrations with the two asymmetric designs, whereas the smooth bootstrap is anticonservative for highly asymmetric designs. Finally, we only observe a slight bias toward overestimating the nominal significance level with random labeling in the simple inhibition point processes. In Table 2 we observe that the smooth bootstrap algorithm performs better than random labeling for Model II. As in the analysis of Model I, we observe some biases toward under or overestimating the nominal significance levels, and we obtain better calibrations for unbalanced designs than for the balanced ones.

Table 1. Probability of rejecting the null hypothesis, p, and 95% confidence interval, $[p_{0.025}, p_{0.975}]$, under $\mathcal{H}_0$ for Model I (Figure 1) at significance level $\alpha = 0.05$.

				Random labeling			Smooth bootstrap		
		bw		df	cv	aicc	df	cv	aicc
IPP	$m_2 = m_2$	11.006	p	0.06	0.062	0.06	0.062	0.06	0.060
			$p_{0.025}$	0.045	0.047	0.045	0.047	0.045	0.045
			$p_{0.975}$	0.075	0.077	0.075	0.077	0.075	0.075
	$m_1 = 2m_2$	11.46	p	0.072	0.072	0.072	0.06	0.064	0.060
			$p_{0.025}$	0.056	0.056	0.056	0.045	0.049	0.045
			$p_{0.975}$	0.088	0.088	0.088	0.075	0.079	0.075
	$m_1 = 9m_2$	11.007	p	0.07	0.076	0.074	0.024	0.022	0.024
			$p_{0.025}$	0.054	0.06	0.058	0.015	0.013	0.015
			$p_{0.975}$	0.086	0.092	0.09	0.033	0.031	0.033
TC	$m_1 = m_2$	11.594	p	0.03	0.032	0.032	0.05	0.052	0.050
			$p_{0.025}$	0.019	0.021	0.021	0.036	0.038	0.036
			$p_{0.975}$	0.041	0.043	0.043	0.064	0.066	0.064
	$m_1 = 2m_2$	11.278	p	0.04	0.04	0.04	0.04	0.04	0.040
			$p_{0.025}$	0.028	0.028	0.028	0.028	0.028	0.028
			$p_{0.975}$	0.052	0.052	0.052	0.052	0.052	0.052
	$m_1 = 9m_2$	11.512	p	0.066	0.066	0.066	0.116	0.112	0.114
			$p_{0.025}$	0.051	0.051	0.051	0.096	0.092	0.094
			$p_{0.975}$	0.081	0.081	0.081	0.136	0.132	0.134
SI	$m_1 = m_2$	11.417	p	0.09	0.105	0.1	0.055	0.055	0.055
			$p_{0.025}$	0.072	0.086	0.081	0.041	0.041	0.041
			$p_{0.975}$	0.108	0.124	0.119	0.069	0.069	0.069
	$m_1 = 2m_2$	11.144	p	0.054	0.054	0.046	0.044	0.044	0.058
			$p_{0.025}$	0.040	0.040	0.033	0.031	0.031	0.044
			$p_{0.975}$	0.068	0.068	0.059	0.057	0.057	0.072
	$m_1 = 9m_2$	11.623	p	0.056	0.062	0.06	0.064	0.064	0.066
			$p_{0.025}$	0.042	0.047	0.045	0.049	0.049	0.051
			$p_{0.975}$	0.07	0.077	0.075	0.079	0.079	0.081

Performance of the no effect test with random labeling and smooth bootstrap calibration in application to Poisson (IPP), Thomas cluster (TC) and simple inhibition (SI) point processes with balanced and unbalanced designs. Calibration procedure and bandwidth selector for the kernel regression function (11) in columns.

In Tables 3 and 4 we analyze the performance of the regression test under the alternative hypothesis considering two different degrees of departure from $\mathcal{H}_0$. As expected, the power of the test increases as differences between the conditional intensities of $\mathbf{X_1}$ and $\mathbf{X_2}$ get larger. In fact, prior to focus on the probabilities of rejection, we can see that the bandwidths of the kernel log-relative risk function are smaller under the alternative hypothesis. The test is able to detect differences between the marginal patterns even in case of conservative calibrations. The performance of the test depends on the type of point process under study, with high powers for clustered point patterns, whereas the probabilities of rejection are low for the simple inhibition point processes, i,e, those with the smoothest conditional intensity functions. Finally, we obtain lower probabilities of rejection when the bandwidth of the kernel regression function is obtained by the approximate degrees of freedom selector.

The results of this simulation study do not allow us to choose one calibration procedure over the other, but random labeling performs slightly better than smooth bootstrap when dealing with highly unbalanced designs.

3.2. Comparison of nonparametric tests

We now apply the three nonparametric tests introduced in Sec. 2 to compare the spatial distribution of Poisson, clustered and regular point processes. The regression test has been implemented as explained in Sec. 3.1. In the Kolmogorov-Smirnov test we have used the two $\pi-$ systems introduced in Sec. 2.1 with $K=6$ and $K=7$. To implement the Cramer von Mises test (Sec. 2.2) we have used plug-in bandwidth selectors in the kernel estimators of the functionals involved in $\hat{T}$

Table 2. Performance of the F-test under the null hypothesis ro the different scenarios of Model II (Figure 2). see details in the caption of Table 1.

				Random labeling			Smooth bootstrap		
		bw		df	cv	aicc	df	cv	aicc
IPP	$m_1 = m_2$	0.559	p	0.078	0.078	0.078	0.038	0.04	0.038
			$p_{0.025}$	0.061	0.061	0.061	0.026	0.028	0.026
			$p_{0.975}$	0.095	0.095	0.095	0.05	0.052	0.05
	$m_1 = 2m_2$	0.57	p	0.034	0.034	0.036	0.03	0.03	0.03
			$p_{0.025}$	0.023	0.023	0.024	0.019	0.019	0.019
			$p_{0.975}$	0.045	0.045	0.048	0.041	0.041	0.041
	$m_1 = 9m_2$	0.576	p	0.012	0.014	0.012	0.044	0.034	0.036
			$p_{0.025}$	0.005	0.007	0.005	0.031	0.023	0.024
			$p_{0.975}$	0.019	0.021	0.019	0.057	0.045	0.048
TC	$m_1 = m_2$	0.556	p	0.016	0.016	0.016	0.102	0.102	0.102
			$p_{0.025}$	0.008	0.008	0.008	0.083	0.083	0.083
			$p_{0.975}$	0.024	0.024	0.024	0.121	0.121	0.121
	$m_1 = 2m_2$	0.551	p	0.038	0.036	0.036	0.086	0.078	0.078
			$p_{0.025}$	0.026	0.024	0.024	0.069	0.061	0.061
			$p_{0.975}$	0.05	0.048	0.048	0.103	0.095	0.095
	$m_1 = 10m_2$	0.551	p	0.064	0.064	0.064	0.062	0.062	0.062
			$p_{0.025}$	0.049	0.049	0.049	0.047	0.047	0.047
			$p_{0.975}$	0.079	0.079	0.079	0.077	0.077	0.077
SI	$B_2 = m_2$	0.553	p	0.022	0.018	0.018	0.026	0.036	0.034
			$p_{0.025}$	0.013	0.01	0.01	0.016	0.024	0.023
			$p_{0.975}$	0.031	0.026	0.026	0.036	0.048	0.045
	$m_1 = 2m_2$	0.542	p	0.044	0.04	0.042	0.076	0.082	0.082
			$p_{0.025}$	0.031	0.028	0.03	0.06	0.065	0.065
			$p_{0.975}$	0.057	0.052	0.054	0.092	0.099	0.099
	$m_1 = 9m_2$	0.551	p	0.048	0.048	0.048	0.036	0.036	0.036
			$p_{0.025}$	0.035	0.035	0.035	0.024	0.024	0.024
			$p_{0.975}$	0.061	0.061	0.061	0.048	0.048	0.048

Table 3. Performance of the regression test under the null and alternative hypotheses for inhomogeneous Poisson (IPP), Thomas cluster (TC) and simple inhibition (Si) point processes with first-order an conditional intensities provided in Figure 1. See details in the caption of Table 1.

				Random labeling			Smooth bootstrap		
			bw	df	cv	aicc	df	cv	aicc
IPP	$\alpha_2 = 2$	$m_1 = m_2$	11.006	0.060	0.062	0.060	0.062	0.060	0.060
		$m_1 = 2m_2$	11.144	0.072	0.072	0.072	0.060	0.064	0.060
		$m_1 = 9m_2$	11.623	0.070	0.076	0.074	0.024	0.022	0.024
	$\alpha_2 = 3$	$m_1 = m_2$	8.229	0.334	0.372	0.368	0.298	0.248	0.272
		$m_1 = 2m_2$	4.683	0.294	0.304	0.294	0.398	0.394	0.400
		$m_1 = 9m_2$	7.442	0.218	0.212	0.212	0.214	0.218	0.204
	$\alpha_2 = 4$	$m_1 = m_2$	5.139	0.748	0.678	0.728	0.776	0.810	0.818
		$m_1 = 2m_2$	8.338	0.808	0.720	0.720	0.774	0.762	0.764
		$m_1 = 9m_2$	7.418	0.314	0.340	0.340	0.348	0.336	0.348
TC	$\alpha_2 = 2$	$m_1 = m_2$	11.594	0.030	0.032	0.032	0.050	0.052	0.050
		$m_1 = 2m_2$	11.278	0.040	0.040	0.040	0.040	0.040	0.040
		$m_1 = 9m_2$	11.512	0.066	0.066	0.066	0.116	0.112	0.114
	$\alpha_2 = 3$	$m_1 = m_2$	0.993	0.908	1.000	1.000	0.838	1.000	1.000
		$m_1 = 2m_2$	0.986	1.000	1.000	1.000	1.000	1.000	1.000
		$m_1 = 9m_2$	0.931	0.482	1.000	1.000	0.264	1.000	1.000
	$a_2 = 4$	$m_1 = m_2$	1.032	0.828	1.000	1.000	0.738	1.000	1.000
		$m_1 = 2m_2$	0.958	0.870	1.000	1.000	0.860	1.000	1.000
		$m_1 = 9m_2$	1.169	1.000	1.000	1.000	0.994	1.000	1.000
SI	$\alpha_2 = 2$	$m_1 = m_2$	11.417	0.090	0.105	0.100	0.055	0.055	0.055
		$m_1 = 2m_2$	11.144	0.054	0.054	0.054	0.046	0.044	0.044
		$m_1 = 9m_2$	11.623	0.056	0.062	0.060	0.064	0.064	0.066
	$\alpha_2 = 3$	$m_1 = m_2$	6.955	0.130	0.080	0.085	0.240	0.275	0.260
		$m_1 = 2m_2$	4.683	0.140	0.396	0.378	0.212	0.540	0.508
		$m_1 = 9m_2$	7.442	0.086	0.112	0.104	0.118	0.228	0.214
	$\alpha_2 = 4$	$m_1 = m_2$	9.024	0.135	0.155	0.155	0.120	0.140	0.140
		$m_1 = 2m_2$	8.338	0.060	0.058	0.050	0.138	0.160	0.148
		$m_1 = 9m_2$	7.418	0.272	0.272	0.270	0.232	0.262	0.268

Table 4. Performance of the regression test under the null and alternative hypotheses for inhomogeneous Poisson (IPP), Thomas cluster (TC) and simple inhibition (Si) point processes with first-order an conditional intensities provided in Figure 2. See details in the caption of Table 1.

				Random labeling			Smooth bootstrap		
			bw	df	cv	aicc	df	cv	aicc
ipp	$\epsilon = 0$	$m_1 = m_2$	0.559	0.078	0.078	0.078	0.038	0.040	0.038
		$m_1 = 2m_2$	0.570	0.034	0.034	0.036	0.030	0.030	0.030
		$m_1 = 9m_2$	0.576	0.012	0.014	0.012	0.044	0.034	0.036
	$\epsilon = 0.4$	$m_1 = m_2$	0.326	0.176	0.216	0.228	0.136	0.130	0.134
		$m_1 = 2m_2$	0.342	0.152	0.196	0.180	0.138	0.150	0.178
		$m_1 = 9m_2$	0.454	0.040	0.046	0.038	0.044	0.040	0.046
	$\epsilon = 0.8$	$m_1 = m_2$	0.148	0.478	0.714	0.866	0.462	0.846	0.894
		$m_1 = 2m_2$	0.146	0.218	0.726	0.834	0.160	0.632	0.766
		$m_1 = 9m_2$	0.279	0.174	0.362	0.320	0.282	0.460	0.458
TC	$\epsilon = 0$	$m_1 = m_2$	0.559	0.078	0.078	0.078	0.038	0.040	0.038
		$m_1 = 2m_2$	0.570	0.034	0.034	0.036	0.030	0.030	0.030
		$m_1 = 9m_2$	0.551	0.064	0.064	0.064	0.062	0.062	0.062
	$\epsilon = 0.4$	$m_1 = m_2$	0.326	0.176	0.216	0.228	0.136	0.130	0.134
		$m_1 = 2m_2$	0.342	0.152	0.196	0.180	0.138	0.150	0.178
		$m_1 = 9m_2$	0.053	1.000	1.000	1.000	0.998	1.000	1.000
	$\epsilon = 0.8$	$m_1 = m_2$	0.148	0.478	0.714	0.866	0.462	0.846	0.894
		$m_1 = 2m_2$	0.146	0.218	0.726	0.834	0.160	0.632	0.766
		$m_1 = 9m_2$	0.051	1.000	1.000	1.000	0.624	0.998	1.000
SI	$\epsilon = 0$	$m_1 = m_2$	0.553	0.022	0.018	0.018	0.026	0.036	0.034
		$m_1 = 2m_2$	0.542	0.044	0.040	0.042	0.076	0.082	0.082
		$m_1 = 9m_2$	0.551	0.048	0.048	0.048	0.036	0.036	0.036
	$\epsilon = 0.4$	$m_1 = m_2$	0.220	0.296	0.576	0.590	0.438	0.694	0.700
		$m_1 = 2m_2$	0.351	0.136	0.172	0.172	0.180	0.210	0.204
		$m_1 = 9m_2$	0.053	0.172	0.200	0.142	0.124	0.184	0.104
	$\epsilon = 0.8$	$m_1 = m_2$	0.135	0.558	0.872	0.906	0.548	0.902	0.922
		$m_1 = 2m_2$	0.359	0.058	0.092	0.086	0.072	0.096	0.092
		$m_1 = 9m_2$	0.051	0.116	0.116	0.092	0.136	0.088	0.038

(Chacón and Duong 2010), and in the kernel intensity estimator used in the bootstrap calibration (Fuentes-Santos, González-Manteiga, and Mateu 2016). The bootstrap calibration was conducted with $B = 200$ realizations of the null hypothesis, as done in the regression test.

Table 5 shows the probability of rejecting $\mathcal{H}_0$ with the three tests for the different scenarios shown in Figure 1. The Kolmvogorov-Smirnov test provides accurate estimators of the nominal significance level for the Poisson (IPP) and regular (SI) balanced and unbalanced simulated patterns with the two $\pi-$ systems and partition sizes, whereas for the clustered point processes the performance depends on K, with a good calibration for $K = 6$, but large probabilities of rejection for $K = 7$. The Cramer von Mises test provides good estimators of the nominal significance level for Poisson and regular point processes, except for highly asymmetric patterns in the first case, and it is slightly anticonservative for clustered point processes. Therefore, under a proper selection of K, the Kolmogorov-Smirnov test is more robust than the two kernel-based tests in terms of calibration. The probability of rejecting $\mathcal{H}_0$ increases as we deviate from the null hypothesis for the three tests, and we observe a faster power increase for kernel-based tests than for the Kolmogorov-Smirnov tests. In fact, this procedure is not able to detect differences between point patterns for some clustered and regular patterns.

Table 6 shows the probability of rejecting $\mathcal{H}_0$ with the three tests for the different scenarios shown in Figure 2. The Kolmvogorov-Smirnov test provides accurate estimators of the nominal significance level for the Poisson (IPP) and regular (SI) balanced and unbalanced pattern with the two $\pi-$ systems and partition sizes, but leads to large type I errors for the clustered patterns. The Cramer von Mises test reports slight deviations from the nominal significance level for the Poisson patterns, it is more anticonservative than the regression test for clustered point patterns, but less sensitive to the sample designs than the regression test. In case of good calibrations, the

Table 5. Performance of the nonparametric tests under the null hypothesis for inhomogeneous (IPP), Thomas cluster (TC) and simple inhibition (Si) with first-order an conditional intensities provided in Figure 1 for balanced ($N_1 = N_2$) and unbalanced ($N_1 = 2N_2=$ designs.

			Kolmovorlog-Smirnov				CvM	No-effect	
			T!a	T!b	T2a	T2b		RL	SB
IPP	$\alpha_2 = 2$	$m_1 = m_2$	0.036	0.036	0.044	0.042	0.048	0.062	0.06
		$m_1 = 2m_2$	0.028	0.02	0.032	0.034	0.052	0.072	0.064
		$m_1 = 9m_2$	0.05	0.06	0.04	0.05	0.014	0.076	0.022
	$\alpha_2 = 3$	$m_1 = m_2$	0.102	0.119	0.118	0.143	0.17	0.372	0.248
		$m_1 = 2m_2$	0.124	0.129	0.13	0.119	0.094	0.304	0.394
		$m_1 = 9m_2$	0.116	0.126	0.096	0.108	0.088	0.212	0.218
	$\alpha_2 = 4$	$m_1 = m_2$	0.34	0.406	0.434	0.467	0.442	0.678	0.81
		$m_1 = 2m_2$	0.399	0.499	0.447	0.489	0.438	0.72	0.762
		$m_1 = 9m_2$	0.258	0.274	0.23	0.252	0.046	0.34	0.336
TC	$\alpha_2 = 2$	$m_1 = m_2$	0.05	0.156	0.052	0.128	0.07	0.032	0.052
		$m_1 = 2m_2$	0.066	0.084	0.048	0.076	0.072	0.04	0.04
		$m_1 = 9m_2$	0.066	0.084	0.048	0.076	0.072	0.066	0.112
	$\alpha_2 = 3$	$m_1 = m_2$	0.364	0.75	0.526	0.82	1	1	1
		$m_1 = 2m_2$	0.642	0.864	0.798	0.93	1	1	1
		$m_1 = 9m_2$	0.552	0.638	0.448	0.672	1	1	1
	$\alpha_2 = 4$	$m_1 = m_2$	0.004	0.052	0	0.002	1	1	1
		$m_1 = 2m_2$	0.92	0.982	0.64	0.84	1	1	1
		$m_1 = 9m_2$	0.92	0.982	0.64	0.84	1	1	1
SI	$\alpha_2 = 2$	$m_1 = m_2$	0.044	0.036	0.042	0.048	0.084	0.105	0.055
		$m_1 = 2m_2$	0.052	0.042	0.052	0.048	0.064	0.054	0.044
		$m_1 = 9m_2$	0.04	0.036	0.048	0.044	0.064	0.062	0.064
	$a_2 = 3$	$m_1 = m_2$	0.132	0.156	0.09	0.118	0.256	0.08	0.275
		$m_1 = 2m_2$	0.042	0.062	0.048	0.082	0.196	0.396	0.54
		$m_1 = 9m_2$	0.024	0.026	0.022	0.014	0.302	0.112	0.228
	$\alpha_2 = 4$	$m_1 = m_2$	0.192	0.224	0.132	0.172	0.286	0.155	0.14
		$m_1 = 2m_2$	0.044	0.038	0.03	0.024	0.34	0.058	0.16
		$m_1 = 9m_2$	0.092	0.096	0.04	0.05	0.44	0.272	0.262

Probability of rejecting $\mathcal{H}_0$ and 95% confidence interval, $[\hat{p}_{0.025}, \hat{p}_{0.975}]$, under $\mathcal{H}_0$ at significance levels $\alpha = 0.05$. In columns $Ks_{i,k}$ denotes the Kolmogorov-Smirnov test, where $i = 1$, 2 denotes the first and second $\pi-$ system, and $K = 6$, 7 the size of the partition. T test denotes the Cramer von Mises test. *df*, *aicc* and *cv* denote the Bandwidth selector for the kernel regression function (11) in the F-test.

power of the kernel-based tests increases as we deviate from the null hypothesis, with faster rates for clustered point processes, i.e., those with the more inhomogeneous conditional intensities. However, the power of the Kolmogorov-Smirnov test is considerably small and in some cases, such as the regular patterns with unbalanced design and $\epsilon = 0.4$, the probability of rejection is even smaller than the type I error under the null hypothesis.

This simulation study shows that the kernel-based tests perform better that the Kolmogorov-Smirnov test, which reported a bad calibration for clustered point processes and was not able to detect differences between regular patterns. Although we cannot choose one kernel-based test over the other, comparison between them suggests using the regression test with unbalanced designs, specially with highly inhomogeneous point patterns such as the clustered point processes in this study, whereas the Cramer von Mises test is more suitable for balanced designs.

4. Application to real data

In this section we compare the performance of the three nonparametric tests through the analysis of wildfire patterns in Galicia (NW Spain) and gunfire in the Rio de Janeiro metropolitan area (Brazil). With these examples we illustrate the utility of nonparametric tests to compare the spatial distribution of point processes in research areas such as environmental risk assessment and criminology.

Table 6. Performance of the nonparametric tests under the null ($a_2 = 2$) and alternative ($a_2 = 3, 4$) hypotheses for inhomogeneos Poisson (IPP), Thomas cluster (TC) and simple inhibition (Si) with first-order an conditional intensities provided in Figure 1.

			Kolmovorlog-Smirnov				CvM	No-effect	
			T!a	T!b	T2a	T2b		RL	SB
IPP	$\epsilon = 0$	$m_1 = m_2$	0.060	0.084	0.062	0.082	0.082	0.078	0.040
		$m_1 = 2m_2$	0.048	0.070	0.040	0.066	0.030	0.034	0.030
		$m_1 = 9m_2$	0.078	0.098	0.040	0.068	0.076	0.014	0.034
	$\epsilon = 0.4$	$m_1 = m_2$	0.090	0.150	0.056	0.092	0.202	0.216	0.130
		$m_1 = 2m_2$	0.066	0.116	0.034	0.070	0.266	0.196	0.150
		$m_1 = 9m_2$	0.040	0.076	0.032	0.060	0.028	0.046	0.040
	$\epsilon = 0.8$	$m_1 = m_2$	0.130	0.296	0.048	0.116	0.876	0.714	0.846
		$m_1 = 2m_2$	0.068	0.200	0.036	0.082	0.724	0.726	0.632
		$m_1 = 9m_2$	0.022	0.108	0.020	0.062	0.462	0.362	0.460
TC	$\epsilon = 0$	$m_1 = m_2$	0.342	0.514	0.242	0.334	0.098	0.078	0.040
		$m_1 = 2m_2$	0.220	0.300	0.150	0.220	0.092	0.034	0.030
		$m_1 = 9m_2$	0.064	0.198	0.032	0.122	0.066	0.064	0.062
	$\epsilon = 0.4$	$m_1 = m_2$	0.304	0.610	0.416	0.642	1.000	0.216	0.130
		$m_1 = 2m_2$	0.010	0.026	0.014	0.030	1.000	0.196	0.150
		$m_1 = 9m_2$	0.328	0.598	0.026	0.090	1.000	1.000	1.000
	$\epsilon = 0.8$	$m_1 = m_2$	0.962	0.998	0.022	0.096	1.000	0.714	0.846
		$m_1 = 2m_2$	0.012	0.034	0.004	0.020	1.000	0.726	0.632
		$m_1 = 9m_2$	0.206	0.480	0.004	0.020	1.000	1.000	0.998
SI	$\epsilon = 0$	$m_1 = m_2$	0.082	0.084	0.064	0.060	0.040	0.018	0.036
		$m_1 = 2m_2$	0.050	0.036	0.050	0.050	0.032	0.040	0.082
		$m_1 = 9m_2$	0.028	0.046	0.042	0.046	0.04	0.048	0.036
	$\epsilon = 0.4$	$m_1 = m_2$	0.15	0.228	0.056	0.092	0.474	0.576	0.694
		$m_1 = 2m_2$	0.052	0.08	0.032	0.046	0.33	0.172	0.21
		$m_1 = 9m_2$	0.074	0.082	0.066	0.06	0.246	0.2	0.184
	$\epsilon = 0.8$	$m_1 = m_2$	0.154	0.268	0.06	0.108	0.844	0.872	0.902
		$m_1 = 2m_2$	0.068	0.1	0.088	0.132	0.372	0.092	0.096
		$m_1 = 9m_2$	0.076	0.094	0.042	0.05	0.122	0.116	0.088

4.1. Comparison of wildfire patterns in Galicia

Wildfire is the most ubiquitous natural disturbance in the world and represents a problem of considerable social and environmental importance. Galicia (NW Spain, Figure 3) has suffered a high incidence of arson fires over the last decades, whereas the risk of forest fires caused by natural factors is low. This problem has motivated an increasing interest in the development of statistical techniques to understand the behavior of wildfires in this region and develop efficient fire prevention and fighting plans (Fuentes-Santos, Marey-Pérez, and González-Manteiga 2013; Ríos-Pena et al. 2015; Boubeta, Lombardía, and Morales 2016).

In this section we consider a dataset comprising the spatial locations and time of occurrence of arson and natural wildfires in Galicia during the period $2004 - 2008$. We have applied the nonparametric tests to compare the spatial distribution of arson and natural fires, and test whether the distribution of arson and natural wildfires varied between years. The Kolmogorov-Smirnov test has been implemented using the second $\pi-$ system introduced in Sec. 2.1, applicable when W is any region in $\mathbb{R}^2$, and we had to reduce the number of partitions in the normalizing constant to $K = 3$ to deal with data sparseness. The Cramer von Mises and regression tests have been applied as in the simulation study with $B = 200$ realizations of the resampling procedures used for calibration. The regression test has been implemented using the three bandwidth selectors in the kernel regression function, $\hat{m}$, but here we show the results obtained with the cross-validation bandwidth because the three procedures provided consistent results.

Figure 4 shows the kernel intensity estimators of the arson and natural wildfires in Galicia for the 5 years under study. As we can see in the top rows of Table 7, we are dealing with highly asymmetric designs. The two kernel-based tests found differences between the spatial distribution

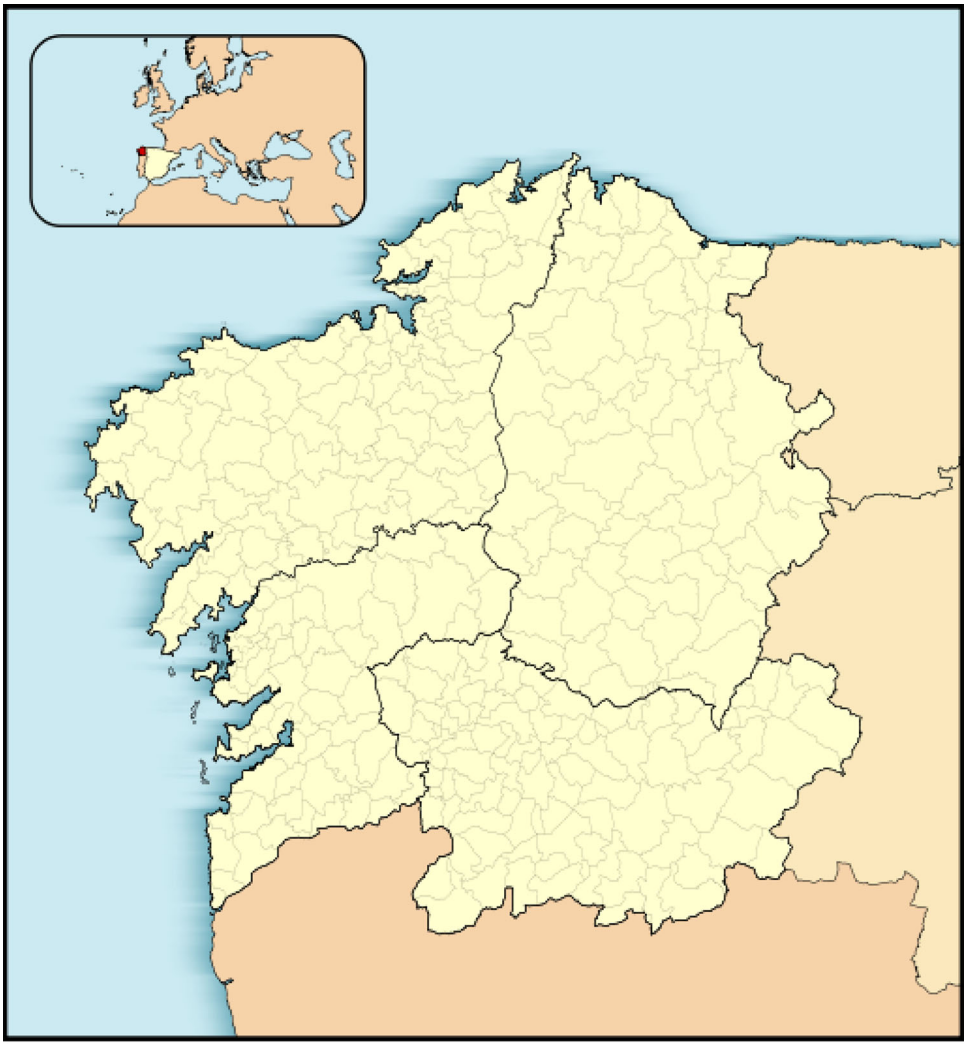

Figure 3. Galicia (NW Spain).

of arson and natural fires for the 5 years under study, whereas the Kolmogorov-Smirnov test did not reject the null hypothesis. The good performance of the kernel-based tests when dealing with unbalanced designs in the simulation study and visual inspection of the kernel intensity estimators in Figure 4 suggest that we should rely on the results provided by the kernel-based tests. In Table 8 we see that the kernel-based tests rejected the null hypothesis for the comparison of arson fires between years, whereas the Kolmogorov-Smirnov test accepted the null in some cases, ($2004-2007$, $2006-2007$ and $2006-2008$). As well as for comparison between arson and natural fires, the intensity estimators in Figure 4 (top) suggest that we should rely on the kernel-based tests. Notice that the Kolmogorov-Smirnov test also reported some difficulties to detect the alternative hypothesis in the simulation study (Sec. 3). Finally, although the intensity estimators in Figure 4 (bottom) suggest that the spatial distribution of natural fires varied over years, the Kolmogorov-Smirnov test did not reject the null hypothesis, whereas the Cramer von Mises and regression tests provided contradictory results (e.g., $2004-2005$) or did not found clear evidences to accept or reject $\mathcal{H}_0$ ($p-$ value $\in (0.01, 0.1)$). These results illustrate the low power and limited performance of this type of procedures when dealing with small-size point patterns.

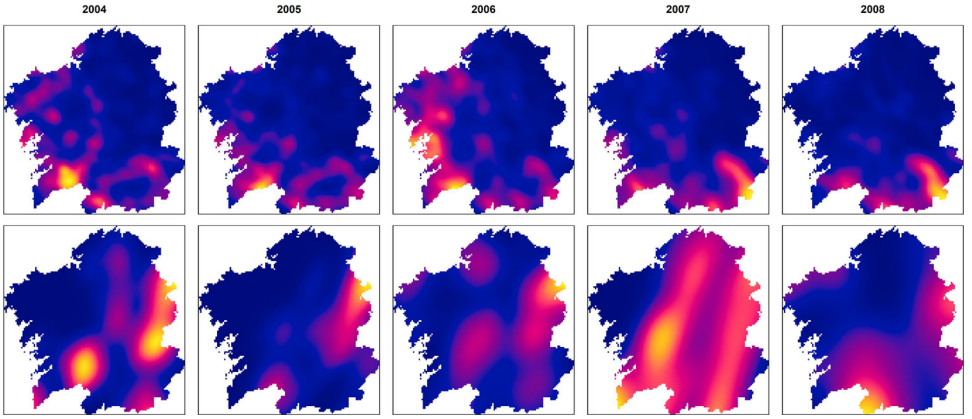

Figure 4. Kernel intensity estimators of arson (top) and natural (bottom) wildfires registered in Galicia in the period 2004-2008 (Figures have different scale).

Table 7. Number of arson and natural fires and p-values of the Kolmogorov-Smirnov test with the second $\pi-$systesms introduced in Sec. 2.1 and $K = 3$.

	2004	2005	2006	2007	2008
Arson	8754	9538	5020	2111	1782
Natural	101	72	119	48	44
KS_2	>0.05	>0.05	>0.05	>0.05	>0.05
CvM	<0.005	<0.005	<0.005	<0.005	<0.005
F-test RL	<0.005	<0.005	<0.005	<0.005	<0.005
F-test SB	<0.005	<0.005	<0.005	<0.005	<0.005

Crameer von Mises test with smooth bootstrap calibration ($B = 200$), and regression test with random labeling (RL) and smooth bootstrap (SB) calibrations ($B = 200$).

Table 8. Nonparametric comparison of the spatial distribution of arson and natural fires across years. See details in the caption of Table 7.

		Arson				Natural			
		2004	2005	2006	2007	2004	2005	2006	2007
2005	KS_1	<0.05				>0.05			
	CvM	<0.005				<0.005			
	F-test RL	<0.005				0.29			
	F-test SB	<0.005				0.355			
2006	KS_1	<0.05	<0.05			>0.05	>0.05		
	CvM	<0.005	<0.005			<0.005	<0.005		
	F-test RL	<0.005	<0.005			0.06	0.01		
	F-test SB	<0.005	<0.005			0.04	0.015		
2007	KS_1	>0.05	<0.05	>0.05		>0.05	>0.05	>0.05	
	CvM	<0.005	<0.005	<0.005		<0.005	0.015	0.015	
	F-test RL	<0.005	<0.005	<0.005		0.015	0.2	0.04	
	F-test SB	<0.005	<0.005	<0.005		<0.005	0.28	0.055	
2008	KS_1	<0.05	<0.05	>0.05	<0.05	>0.05	>0.05	>0.05	>0.05
	CvM	<0.005	<0.005	<0.005	<0.005	0.195	0.015	0.065	0.08
	F-test RL	<0.005	<0.005	<0.005	<0.005	0.08	0.34	0.005	0.385
	F-test SB	<0.005	<0.005	<0.005	<0.005	0.06	0.34	0.02	0.225

4.2. Comparison of gunfire patterns in Rio De Janeiro metropolitan area

As police departments have created centralized databases of crime reports comprising, among other information, the location and time of occurrence of each event, point process modeling has been widely used to predict the risk of future crime. In the particular case of the Rio de Janeiro

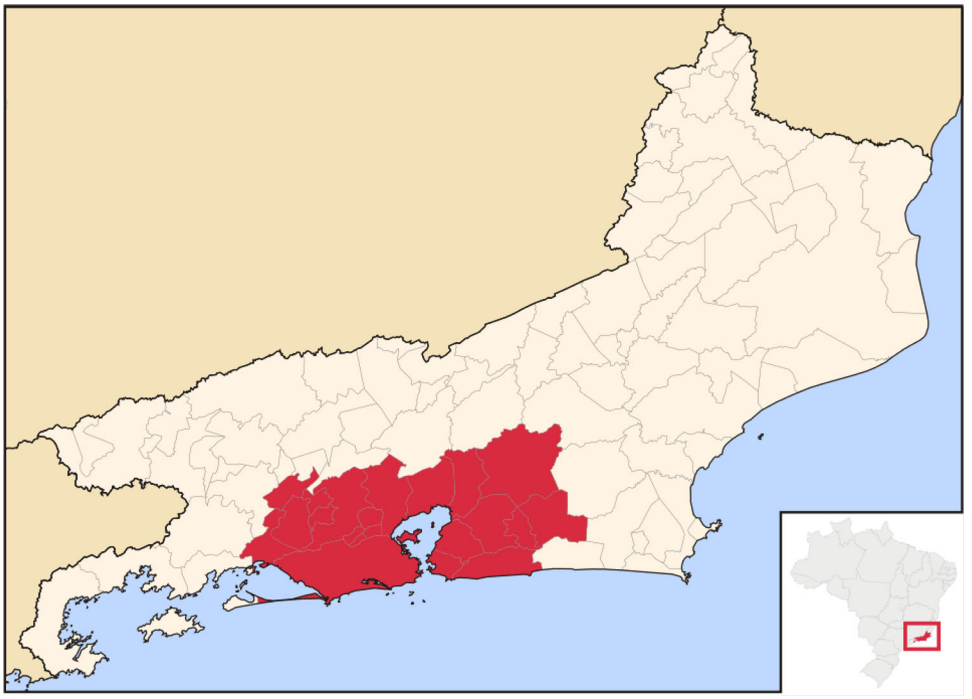

Figure 5. Rio de Janeiro metropolitan area (Brazil).

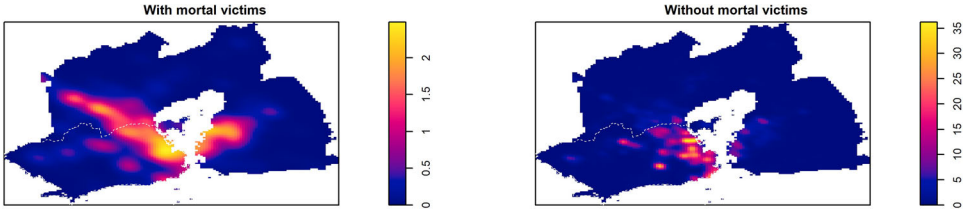

Figure 6. Kernel intensity function for the spatial patterns of gunfire events with and without mortal victims (Figures have different scale).

metropolitan area (Brazil, Figure 5), which has been suffering a continuous increase of gun violence over the last decades (Arias and Barnes 2017), the ISP-RJ centralizes reports from the 190 crime report hotline, military, and civil police in a global crime database. In parallel with the official sources, the collaborative mobile app *Fogo Cruzado*, which collects real time gunfire reports and delivers instant alerts to help citizens avoid stray bullets, has generated a valuable data set of gunfire violence in Rio de Janeiro.

Here we have applied the three nonparametric tests to compare the spatial distribution of the 1141 gunfire events with mortal victims and the 4804 events without victims collected by *Fogo Cruzado* in the Rio de Janeiro metropolitan area during 2017 (Fuentes-Santos, González-Manteiga, and Zubelli 2020). We used the second $\pi-$ system introduced in Sec. 2.1 for the Kolmogover-Smirnov test and $K=4$ partitions in the normalizing constant to deal with data sparseness. As above, the bootstrap calibrations in the kernel-based tests were ran with $B=200$ realizations of the null hypothesis. For the regression test we report the results obtained with the cross-validation bandwidth in $\hat{m}$ as the three procedures provided similar results.

The Kolmogorov-Smirnov test did not detect differences between gunfire patterns (p-value $>$ 0.05), whereas the two kernel-based tests rejected the null hypothesis (p-value $<$ 0.005). Visual

inspection of the kernel intensity estimators (Figure 6) suggests that we should trust on the Cramer von Mises and regression tests. This example confirms the limited capacity of the Kolmogorov-Smirnov test when dealing with highly inhomogeneous and sparse data. The small number of partitions, K, required to deal with data sparseness has an oversmoothing effect on the test statistic and reduces the power of the test.

5. Discussion

The natural way to compare the spatial distribution of two spatial point pattterns is testing whether these patterns have the same first-order structure. Two works have recently proposed nonparametric tests for this purpose. Zhang and Zhuang (2017) proposed a Kolmogorov-Smirnov test using the absolute difference between the point densities of the observed patterns over a $\pi-$ system as discrepancy measure. Fuentes-Santos, González-Manteiga, and Mateu (2017) use a nonparametric test based on an L_2 discrepancy between densities of event locations. Here, we propose a new nonparametric test based on the log-relative risk function, widely used in epidemiology to analyze the spatial distribution of a given disease Bithell (1990); Davies, Jones, and Hazelton (2016); Kelsall and Diggle (1995). Taking into account that the relative risk function of two spatial point processes with the same spatial distribution is constant, we introduce a no-effect test that checks whether the log-relative risk function depends on event locations. This proposal is in line with the log-ratio based separability test proposed by Fuentes-Santos, González-Manteiga, and Mateu (2018) and based on the nonparametric regression test introduced by Bowman and Azzalini (1997). In this work we have compared the performance of the three tests through a simulation study and through application to the analysis of two relevant real data problems.

Prior to analyze the performance of the tests, let us compare their formulation and implementations. Here we focus on three key aspects: (i) the implementation of the test, (ii) regularity assumptions and (iii) computational demand. The Kolmogorov-Smirnov test does not require the estimation of first-order properties, avoiding the bandwidth selection problem of nonparametric estimators. However, it should be noted that this test involves the selection of a $\pi-$ system and a number of partitions that can be seen as a smoothing parameter. In contrast, to implement the Cramer von Mises and regression tests we need nonparametric estimators of some first-order properties. The first-order intensities and functionals involved in the L_2 test have been estimated through kernel smoothing with plug-in bandwidths (Chacón and Duong 2010; Fuentes-Santos, González-Manteiga, and Mateu 2016). The log-relative risk function in the no-effect test has also been estimated by kernel smoothing with scalar cross-validation bandwidth (Kelsall and Diggle 1995). The consistency of the kernel estimators used in both tests has been proved for Poisson point processes, a restrictive and unrealistic condition that may be seen as a weakness in comparison with the Kolmogorov-Smirnov test, which allows dependence between events. However, these tests can also be applied to non-Poisson point processes, in which case we compare their spatial distribution measuring the discrepancy between their conditional intensities. Finally, the Kolmogorov-Smirnov test is calibrated through the asymptotic distribution of the null hypothesis and, consequently, computationally efficient, in contrast with the high computational demand of the resampling algorithms required to calibrate the kernel-based tests.

The results of the simulation study indicate that the three nonparametric tests provide reasonable estimators of the nominal significance level under the null hypothesis, although with some deviations from the nominal significance level. Comparison between calibration procedures in the regression test suggest that the random labeling algorithm performs better for unbalanced designs, whereas the bootstrap calibration provides more accurate calibrations for balanced designs. The kernel-based tests perform better than the Kolmogorov-Smirnov procedure in terms of power. In fact, the later was not able to detect differences between patterns in some cases. We also observe a faster power increase for Poisson and clustered point processes, whereas

probabilities of rejection for regular point processes are relatively small in both kernel-based tests. These differences between clustered and regular point processes are linked with the inhomogeneity of the conditional intensities. Kernel estimators of point processes with smooth spatial distribution have larger bandwidth parameters than those with highly inhomogeneous distributions; this higher smoothness in the kernel intensity or log-relative risk function estimators may hamper the detection of differences between the observed patterns. This problem may be solved using adaptive instead of fixed bandwidth parameters, as done by Davies, Jones, and Hazelton (2016). However, the use of adaptive bandwidths may add a considerable computational cost to these tests limiting their applicability to real data analysis.

The application to the analysis of real data shows that the Kolmogorov-Smirnov test has some limitations when dealing with sparse data, as this procedure does not detect the observed differences between arson and natural wildfire patterns. The results provided by the Cramer von Mises and regression tests agree with the visual inspection of the kernel intensity estimators. However, these tests may have a poor performance when dealing with small datasets. In view of these results, we consider that the kernel-based tests could be more accurate than the Kolmogorov-Smirnov tests, particularly in the analysis of spatial point processes with sparse events. From a practical viewpoint, the differences found between arson and natural wildfires in Galicia indicate that we should use different covariates or models to predict the occurrence of each type of wildfire. Similar conclusions can be derived from the analysis of gunfire with and without fatalities in Rio de Janeiro.

6. Conclusions

This work compares the performance of three nonparametric formal tests developed to check if two spatial point patterns have the same spatial distribution, a computationally efficient Kolmogorov-Smirnov test with asymptotic calibration and two computationally demanding kernel-based tests with bootstrap calibration. Kernel-based tests outperform the Kolmogorov-Smirnov test, which reported a poor performance when dealing with sparse data and highly inhomogeneous patterns. Comparison between the kernel-based tests suggests using the Cramer von Mises test to compare spatial point processes with similar number of events, whereas the regression test with random labeling calibration is more suitable for unbalanced designs. However, even the kernel-based tests have a poor performance when dealing with regular patterns and small sample sizes, in both cases derived from the large bandwidths used in the kernel estimators involved in these tests.

Finally, when these tests find differences between two spatial point processes, gunfire with and without mortal victims, for instance, we may wonder where did those differences occur. The kernel log-relative risk functions with tolerance contour surfaces can be used as local test to answer this question, allowing us to detect areas with high mortality risk in the gunfire case study. Here, as we are interested in differences at local scale, we should use kernel estimators with adaptive bandwidth parameters, as proposed by Davies, Jones, and Hazelton (2016). However, acceptance of the null hypothesis does not imply that the observed patterns have been generated by the same point process. Notice that, although the conditional intensity determines the spatial distribution of events in a spatial point process, the same conditional intensity can be generated by different models, a Poisson and a cluster process for instance. Therefore, further information about the dependence structure of the observed patterns and procedures for the combined analysis of first and second-order properties are required to test whether two observed patterns are realizations of the same point process.

Acknowledgements

We would like to acknowledge the collaboration of Prof. Manuel Marey for involving us in the analysis of wildfire in Galicia, and Luis Franco for his support in the access and pre-processing of wildfire data. We are grateful to

Prof. Jorge P. Zubelli for involving us in the analysis of gun violence in Rio de Janeiro, Cecilia Oliveira for providing the gunfire database and for a number of highly clarifying discussions, and Eng. Luiz Claudio Dantas (LAMCA) for technical support in the pre-procession of the gunfire data. We acknowledge the anonymous reviewers and the Associate Editor for their valuable comments that substantially improved this paper.

Funding

This work has been supported by Projects MTM2016-78917-R and MTM2016-76969-P (AEI/FEDER, UE), grant AICO/2019/198 from Generalitat Valenciana, and IAP network StUDyS grant 3E120297 from the Belgian Science Policy.

ORCID

I. Fuentes-Santos http://orcid.org/0000-0002-3298-9648
W. González-Manteiga http://orcid.org/0000-0002-3555-4623

References

Alba-Fernández, M., F. Ariza-López, M. D. Jiménez-Gamero, and J. Rodríguez-Avi. 2016. On the similarity analysis of spatial patterns. *Spatial Statistics* 18:352–62. doi:10.1016/j.spasta.2016.07.004.

Andresen, M. A. 2009. Testing for similarity in area-based spatial patterns: A nonparametric Monte Carlo approach. *Applied Geography* 29 (3):333–45. doi:10.1016/j.apgeog.2008.12.004.

Arias, E. D., and N. Barnes. 2017. Crime and plural orders in Rio de Janeiro. *Current Sociology* 65 (3):448–65. doi: 10.1177/0011392116667165.

Baddeley, A., and R. Turner. 2005. Spatstat: An R package for analyzing spatial point patterns. *Journal of Statistical Software* 12 (6):1–42. www.jstatsoft.org. doi:10.18637/jss.v012.i06.

Bithell, J. 1990. An application of density estimation to geographical epidemiology. *Statistics in Medicine* 9 (6): 691–701. doi:10.1002/sim.4780090616.

Boubeta, M., M. J. Lombardía, and D. Morales. 2016. Empirical best prediction under area-level poisson mixed models. *Test* 25 (3):548–69. doi:10.1007/s11749-015-0469-8.

Bowman, A. W., and A. Azzalini. 1997. *Applied smoothing techniques for data analysis: The kernel approach with S-plus illustrations*. New York: Oxford Statistical Science Series 18.

Bowman, A. W., and A. Azzalini. 2014. *R package sm: Nonparametric smoothing methods (version 2.2-5.5)*. Glasgow: University of Glasgow. http://www.stats.gla.ac.uk/adrian/sm/.

Chacón, J. E., and T. Duong. 2010. Multivariate plug-in bandwidth selection with unconstrained pilot bandwidth matrices. *Test* 19 (2):375–98. doi:10.1007/s11749-009-0168-4.

Cucala, L. 2006. *Espacements bidimensionnels et données entachées ďerreurs dans ľanalyse des processus ponctuels spatiaux*. Ph.D. thesis, Université des Sciences Sociales, Toulouse I.

Davies, T. M., and M. L. Hazelton. 2010. Adaptive kernel estimation of spatial relative risk. *Statistics in Medicine* 29 (23):2423–37. doi:10.1002/sim.3995.

Davies, T. M., K. Jones, and M. L. Hazelton. 2016. Symmetric adaptive smoothing regimens for estimation of the spatial relative risk function. *Computational Statistics & Data Analysis* 101:12–28. doi:10.1016/j.csda.2016.02.008.

Davies, T. M., J. C. Marshall, and M. L. Hazelton. 2018. Tutorial on kernel estimation of continuous spatial and spatiotemporal relative risk. *Statistics in Medicine* 37 (7):1191–221. doi:10.1002/sim.7577.

Diggle, P. J. 2013. *Statistical analysis of spatial and spatio-temporal point patterns*. 3rd ed. Boca Raton, FL: CRC Press.

Duong, T., B. Goud, and K. Schauer. 2012. Closed-form density-based framework for automatic detection of cellular morphology changes. *Proceedings of the National Academy of Sciences of the United States of America* 109 (22):8382–87. doi:10.1073/pnas.1117796109.

Duong, T. 2013. ks: Kernel smoothing. http://CRAN.R-project.org/package=ks.

Fibich, P., J. Lepš, V. Novotný, P. Klimeš, J. Těšitel, K. Molem, K. Damas, and G. D. Weiblen. 2016. Spatial patterns of tree species distribution in new guinea primary and secondary lowland rain forest. *Journal of Vegetation Science* 27 (2):328–39. doi:10.1111/jvs.12363.

Fuentes-Santos, I., W. González-Manteiga, and J. Mateu. 2016. Consistent smooth bootstrap kernel intensity estimation for inhomogeneous spatial poisson point processes. *Scandinavian Journal of Statistics* 43 (2):416–35. doi: 10.1111/sjos.12183.

Fuentes-Santos, I., W. González-Manteiga, and J. Mateu. 2017. A nonparametric test for the comparison of first-order structures of spatial point processes. *Spatial Statistics* 22:240–60. doi:10.1016/j.spasta.2017.02.007.
Fuentes-Santos, I., W. González-Manteiga, and J. Mateu. 2018. A first-order, ratio-based nonparametric separability test for spatiotemporal point processes. *Environmetrics* 29 (1):e2482. doi:10.1002/env.2482.
Fuentes-Santos, I., W. González-Manteiga, and J. Zubelli. 2020. Nonparametric spatiotemporal analysis of violent crime. A case study in the Rio de Janeiro metropolitan area. *Spatial Statistics*. Advance online publication. doi:10.1016/j.spasta.2020.100431.
Fuentes-Santos, I., M. Marey-Pérez, and W. González-Manteiga. 2013. Forest fire spatial pattern analysis in Galicia (NW Spain). *Journal of Environmental Management* 128:30–42. doi:10.1016/j.jenvman.2013.04.020.
Getzin, S., T. Wiegand, K. Wiegand, and F. He. 2008. Heterogeneity influences spatial patterns and demographics in forest stands. *Journal of Ecology* 96 (4):807–20. doi:10.1111/j.1365-2745.2008.01377.x.
Hahn, U. 2012. A studentized permutation test for the comparison of spatial point patterns. *Journal of the American Statistical Association* 107 (498):754–64. doi:10.1080/01621459.2012.688463.
Hering, A. S., C. L. Bell, and M. G. Genton. 2009. Modeling spatio-temporal wildfire ignition point patterns. *Environmental and Ecological Statistics* 16 (2):225–50. doi:10.1007/s10651-007-0080-6.
Illian, D. J., P. A. Penttinen, D. H. Stoyan, and D. Stoyan. 2008. *Statistical analysis and modelling of spatial point patterns*. Chichester: John Wiley & Sons, April.
Juan, P., J. Mateu, and M. Saez. 2012. Pinpointing spatio-temporal interactions in wildfire patterns. *Stochastic Environmental Research and Risk Assessment* 26 (8):1131–50. doi:10.1007/s00477-012-0568-y.
Kelsall, J. E., and P. J. Diggle. 1995. Kernel estimation of relative risk. *Bernoulli* 1 (1–2):3–16. doi:10.2307/3318678.
Nadaraya, E. A. 1964. On estimating regression. *Theory of Probability & Its Applications* 9 (1):141–42. doi:10.1137/1109020.
Perry, G. L., B. P. Miller, and N. J. Enright. 2006. A comparison of methods for the statistical analysis of spatial point patterns in plant ecology. *Plant Ecology* 187 (1):59–82. doi:10.1007/s11258-006-9133-4.
R Core Team. 2019. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
Ríos-Pena, L., C. Cadarso-Suárez, T. Kneib, and M. Marey-Pérez. 2015. Applying Binary Structured Adiditive Regression (STAR) for predicting wildfire in Galicia. *Spain. Procedia Environmental Sciences* 27:123–26.
Ripley, B. 1977. Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society, Series B* 39: 172–212.
Ripley, B. 1981. *Spatial statistics*. New York: Wiley.
Watson, G. S. 1964. Smooth regression analysis. *Sankhya: The Indian Journal of Statistics, Series A* 26 (4):359–72.
Zhang, T., and R. Zhuang. 2017. Testing proportionality between the first-order intensity functions of spatial point processes. *Journal of Multivariate Analysis* 155:72–82. doi:10.1016/j.jmva.2016.11.013.