Nonparametric spatiotemporal analysis of violent crime. A case study in the Rio de Janeiro metropolitan area

Isabel Fuentes-Santos *
Marine Research Institute
Spanish National Research Council

Wenceslao González-Manteiga
Department of Statistics, Mathematical Analysis and Optimization
University of Santiago de Compostela

and
Jorge P. Zubelli
Instituto Nacional de Matemática Pura e Aplicada (IMPA)

December 9, 2019

*We thank Mrs. Cecilia Oliveira for providing the database and for a number of highly clarifying discussions, Eng. Luiz Claudio Dantas (LAMCA) for technical support with data pretreatment and help with auxiliary Rio de Janeiro georeferenced information. This work has been conducted during the visits of Prof. González-Manteiga to the IMPA in April 2018 and March 2019, and the visit of Prof. Zubelli to the University of Santiago de Compostela in October 2018. Dr. Fuentes-Santos and Prof. González-Manteiga have been funded by project MTM2016-76969-P (AEI/FEDER, UE), Prof. Zubelli has been supported by the CNPq under grant 307873 and by FAPERJ through the program Cientistas do Nosso Estado.

Abstract

The collaborative mobile app *Fogo Cruzado* delivers instant alerts every time a user reports a gunfire in the Rio de Janeiro metropolitan area (Brazil). This app contributes to public safety and generates a valuable dataset comprising the location and time of occurrence of gun shootings, which analysis may allow reserchers to understand gunfire dynamics in Rio and support the development of crime reduction plans. Prior to apply existing crime forecasting methods, such as kernel hotspot maps and self-exciting, we should test if the gunfire patterns meet their assumptions. For this purpose, we have applied nonparametric first and second-order point process inference. The kernel intensity estimator describes the spatial distribution of gunfire and identifies chronic hotspots. The nonparametric test for comparison of first-order intensities found differences between gunfires with and without fatalities or police intervention. The recently developed log-ratio based first-order separability test found that the spatial distribution of gunfire, fatalities and police presence varied over time. Finally, spatiotemporal inhomogeneous K-tests detected clustering between gunfire events, fatalities and police interventions. These results suggest that we could consider a self-exciting point process with nonseparable background component as a starting point in the development of a suitable approach to forecast gunfire hotspots in Rio de Janeiro.

Keywords: First-order intensity, *Fogo cruzado*, gunfire, kernel smoothing, inhomogeneous K-function, separability test

1 Introduction

Crime, as many socio-economic and environmental processes such as forest fires, earthquakes or disease outbreaks, takes the form of events occurring irregularly in space and time. Spatial and spatiotemporal point processes, which govern the location of a random number of events in a continuous domain (Diggle, 2013), are a suitable framework for the analysis of this type of data. Point process modeling has been increasingly applied during the last decade to analyze a wide variety of crimes against both property, such as residence or car burglaries (Chainey et al., 2008; Mohler et al., 2011), and citizens, such as homicides and violent crimes (Mohler, 2014; Taddy, 2010).

Point process modeling of crime data has mainly focused on the detection and forecasting of hotspots, i.e, geographic locations of high crime concentration in comparison with the distribution of crime across the whole region of interest (Chainey et al., 2008). We can find two types of crime hotspots, chronic and temporary, which have different nature and require different strategies to be reduced (Mohler, 2014). Chronic or long-term hotspots are characterized by a high crime volume over several years, and need problem oriented policing strategies that focus on the root causes of crime. Temporary or short-term hotspots last on the time scale of days or weeks, may be caused by contagious like processes and can be reduced through a temporary increase of police presence in the affected area. Hotspot prediction models should be able to distinguish between chronic and temporary hotspots in order to aid law enforcement and police to implement accurate crime reduction strategies for each type of hotspot.

The statistical techniques used in crime hotspot forecasting can be grouped into two broad categories, multivariate analysis and event-based nonparametric methods. Multivariate models incorporate additional variables such as demographics (Wang et al., 2012), income levels (Liu and Brown, 2003), distance to crime attractors (Kennedy et al., 2011; Liu and Brown, 2003; Wang et al., 2012), and leading-indicator crimes (Cohen et al., 2007; Gorr, 2009). Static variables such as demographics and distance to crime attractors are used to predict long-term crime hotspots, whereas recent crime activity is useful in short-term hotspot forecasting.

The most popular nonparametric event-based tool in the analysis of crime data is kernel density estimation, which produces static maps of crime occurrence. Chainey et al. (2008) showed the suitability of spatial kernel density estimations for hotspot mapping and prediction of future crime patterns through its application to several types of crime in London, including residential burglaries, street crimes, thieves from vehicles and thieves of vehicles. Gerber (2014) incorporated twitter information as covariates in the kernel density estimator to forecast the occurrence of 25 types of crime in Chicago. Both works discussed bandwidth selection, which is crucial for the good performance of any kernel estimator, Chainey et al. (2008) proposed using a rule of thumbs or a subjective scalar bandwidth, whereas Gerber (2014) used a plug-in bandwidth matrix. However, none of them applied edge-correction to the kernel hotspot maps, which is required to reduce the bias near the boundary when the observation domain is bounded. The main drawback of kernel density estimation and other event-based approaches is that these models are not able to reflect multiple timescales simultaneously. Therefore researchers should choose between long-term or short-term hotspot maps. Long-term maps are generated using several years of data (Weisburd et al., 2012), whereas short-term hotspot maps are based on some weeks or months of data and use spatial bandwidths in the order of tens or hundreds of meters (Chainey et al., 2008; Kennedy et al., 2011; Wang et al., 2012).

Prospective hotspot maps (Bowers et al., 2004), which use spatiotemporal kernel density estimators that give larger weights to recent events than to events further in the past, are a natural extension of static hotspot maps. Research on near-repeat offender behavior (Bowers et al., 2004; Farrell and Pease, 2001; Johnson et al., 2007; Johnson, 2008) suggests that prospective hotspot maps can help police departments in the reduction of contagious effects associated with certain crimes. However, the accuracy of both short-term and prospective hotsptot maps can be reduced as small sample sizes and small bandwidths increase the variance of kernel estimators. Furthermore, prospective maps allow us to estimate chronic or near-repeat hotspots through a suitable selection of the temporal observation domain and bandwidth parameter, but we cannot either estimate both types of hotspots simultaneously nor discriminate between them.

Certain types of crime, such as burglaries and gang violence, present highly clustered patterns. This behavior is close to that observed in seismology, where the occurrence of an earthquake increases the risk of aftershocks in its neighborhood. Considering this similarity and the limitations of both spatial and spatiotemporal kernel hotspot maps, Mohler et al. (2011) proposed using self-exciting point processes, which were developed to model earthquake patterns, in crime modeling and illustrated the suitability of this approach through application to the analysis of residential burglaries in Los Angeles. Self-exciting point processes decompose the rate of crime into a stationary background component, which estimates chronic hotpots, and a component analogous to prospective hotspot maps that captures near-repeat effects. Mohler (2014) extended this model to the framework of marked point processes to predict homicides in Chicago incorporating information of precursory gun crimes to the second component. Reinhart and Greenhouse (2018) incorporated spatial covariates to the background component. Zhuang and Mateu (2018) introduced daily and weekly periodicity in the background component of self-exciting point process models for the analysis of robbery related violence in Castellón, (Spain).

Crime modeling with kernel hotspot maps or self-exciting point processes requires the assumption of a series of hypothesis that may not be fulfilled by the observed crime patterns under study. The spatial kernel estimator assumes that the time of occurrence of crimes is a realization of a stationary Poisson point pattern, whereas prospective maps allow the temporal point process to be nonstationary but Poisson. In both cases the spatial component of the spatiotemporal point process is assumed to be Poisson, i.e., crimes occur independently in both space and time. Self-exciting point process models assume a contagious behavior in crime. As pointed out in recent works (Loeffler and Flaxman, 2018; Wooditch and Weisburd, 2016), these hypothesis need to be tested prior to the application of any hotspot model. Wooditch and Weisburd (2016) applied a spatiotemporal stationary K-test to check whether stop-question-frisk (SQF) practices contribute to reduce crime incidence in New York. Loeffler and Flaxman (2018) combined Bayesian spatiotemporal point process modeling and classical space/time interaction tests to check whether spatiotemporal distinguish between hetereogeneity, referred as endemic or first-order clustering, and clustering associated to epidemic or contagions-like processes in gun violence.

This work analyzes, for the first time, gunfire reports collected by the collaborative app

Fogo Cruzado in the Rio de Janeiro metropolitan area during 2017. *Fogo Cruzado* records the GPS location and time of occurrence of each gunfire event, as well as some additional information, such as number of victims or an indicator of police presence. Therefore we have a realization of a spatiotemporal marked point process that could be modeled through any of the methods outlined above. However, as a first approach to these data we need to conduct a complete exploratory analysis in order to understand the behavior of gun violence in Rio de Janeiro and support the selection of an accurate model to predict future gunfire hotspots. For this purpose we apply novel nonparametric inference techniques for spatiotemporal point processes to analyze the spatial and spatiotemporal distribution of gun shootings and to test for interaction between them. We describe the spatial distribution of gun shootings by kernel intensity estimation, and apply a nonparametric comparison of first-order intensities to test if police presence has the same distribution as gunfire events and to compare the spatial distribution of events with and without victims. A nonparametric separability test allows us to check whether the spatial distribution of gunfire varies over time. Finally, the inhomogeneous spatial and spatiotemporal K-tests check whether the occurrence of a gun shooting in a given location may increase or reduce future crime risk in its neighborhood.

The remainder of this work is organized as follows, in Section 2 we introduce the study area and crime data. Section 3 provides an overview of the nonparametric inference techniques used in this work. In Section 4 we show the results of the spatial and spatiotemporal analysis of the *Fogo Cruzado* data set. We discuss these results and future research lines in Section 5. In the Supporting Information we provide more detailed information for Section 3 for those with are not familiar with these techniques (Appendix A), and the kernel spatiotemporal intensity functions of the gunfire pattrns under study (Appendix B).

2 Data

2.1 Study area and data collection

Rio de Janeiro, the capital of the state of Rio de Janeiro, with a population of 6.45 million in 2017, is the second-most populous municipality in Brazil and the sixth-most populous in the Americas. It is also the largest municipality of the Rio de Janeiro metropolitan area (Figure 1), which has a total population of 12.3 million. Both the city and metropolitan area have suffered a continuous increase in violent crimes during the last decades. This violence increase has been linked to the large social and financial inequalities and to the regional distribution of the population (Arias and Barnes, 2017; Silva et al., 2017). In particular Arias and Barnes (2017), which highlight the need of local-level research on organized crime to understand violent crime dynamics, describe two main types of organized crime in the city, drug gangs in the *Zona Norte* (North zone) and police-linked extortion rackets known as milícias in the *Zona Oeste* (West Zone) of Rio de Janeiro.

Reducing crime rates has become a major concern for the government and public safety institutions (Silva et al., 2017), which have adopted both confrontation and community oriented strategies to fight violent crime. Among the latter we find the UPP (Unidades de Policia Pacificadora) project, which have occupied favelas with high crime rates since 2008 and conducted community oriented methods to reduce violence (Oliveira, 2018; Passos, 2018)

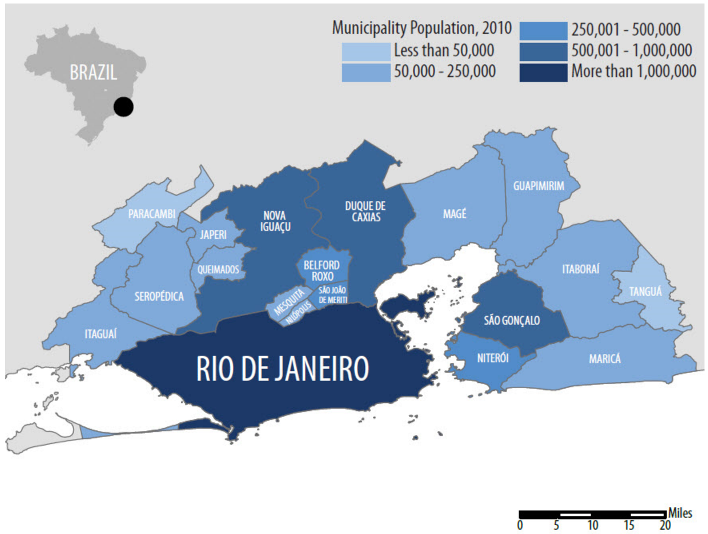

Figure 1: Location and municipalities of the Rio de Janeiro metropolitan area in Brasil. Population from the Census 2010 (IBGE).

The Instituto de Segurança Pública de Rio de janeiro (ISP-RJ) has developed a centralized database including information from the 190 crime report hotline, military, and civil police in the Rio de Janeiro metropolitan area during the last decades. These data have been used in recent studies to develop interactive visualization tools and crime risk maps such as CrimeVis (Silva et al., 2017) and CrimeRadar (Igarapé Institute, `https://igarape.org.br/en/`). Both the centralized database and the crime visualization tools provide support to researchers and public policy makers addressing the challenging task of reducing violent crime in Rio de Janeiro. However, a group of independent researchers found serious difficulties to obtain accurate information about gunfire violence in the press, social media, and police reports when they attempted to quantify gun shootings after an extremely violent outbreak in 2015. These difficulties pointed out the need of a highly visible channel to report gunfire events and motivated the development of *Fogo Cruzado* (`https://fogocruzado.org.br/`), a mobile app that was released for Android (R) and Apple (R) mobile smart phones in June 2016 by Amnesty International. In January 2018 the project became autonomous and independent of Amnesty International, and has been managed by the Update Institute since then. The users of *Fogo Cruzado*, who can identify themselves or not, report gun shootings in real time to the app, which delivers an instant alert to nearby users. In addition to the location, which is obtained through the GPS of the device, the witness can include some details in the report, such as number of victims or whether police was involved in the shooting. The *Fogo cruzado* management team applies a data validation procedure that discards incomplete information, repeated reports, and data that can not be confirmed by the project team. The discard rate from July 2016, when the app was released, to the end of 2018 was 16% of received notifications. The database is completed through the incorporation of information provided by in-site partners, such as collectives, communicators and active residents, as well as with press and law enforcement reports. Therefore, in addition to its contribution public safety helping residents avoid stray bullets, *Fogo Cruzado* generates a valuable database which analysis shall allow researchers to understand the dynamics of violent crime, identify the causes of gun violence and develop more efficient strategies to fight against gun violence.

2.2 Data description

This work analyses the 5945 gun shootings, an average of 16 events per day, reported to *Fogo Cruzado* in the Rio de Janeiro metropolitan area during 2017. Each event is characterized by the spatial coordinates of its location, date and time of occurrence, an indicator of police interventions, and the number of mortal and injured police and civil victims. This work focuses on factor variables that indicate whether gunfire events caused injuries or mortal victims (see details in Table 1 and Figure 2).

Table 1: Gunfire reports in the Rio de Janeiro metropolitan area

	Total	Rio de Janeiro	Metropolitan area
Crimes	5945	3965	1980
Police interventions	686	483	203
With mortal victims	1141	488	653
With injured victims	967	615	352

Fogo Cruzado project published periodic reports summarizing the information collected by the app (`https://fogocruzado.org.br/relatorios-rj/`). The 2017 Annual Balance provides detailed information about the neighborhoods and municipalities with more gunfire reports, rates of fatalities and injuries, police presence during shootings, periods with higher gunfire incidence and higher number of victims. The balance also analyzes gunfire incidence in areas with UPPs, shopping malls, and closure of schools/suspension of classes due to gun violence. Rio de Janeiro (3997 events and 669 mortal victims), São Gonçalo (589 shootings and 232 fatalities) and Niterói (311 events and 59 mortal victims) were the most affected municipalities. In the city of Rio de Janeiro, which suffered the 67% of gunfire events recorded in the whole metropolitan area, the districts of Cidade de Deus and Complexo do Alemão, with 175 notifications each, reported the highest gunfire incidence. The Alemão UPP, which covers an area larger than the favela of Complexo do Alemão, had the highest shooting rate with 193 notifications. Furthermore, 20 gunfire events, with 5 mortal victims, were recorded in shopping malls located in the Rio de Janeiro metropolitan

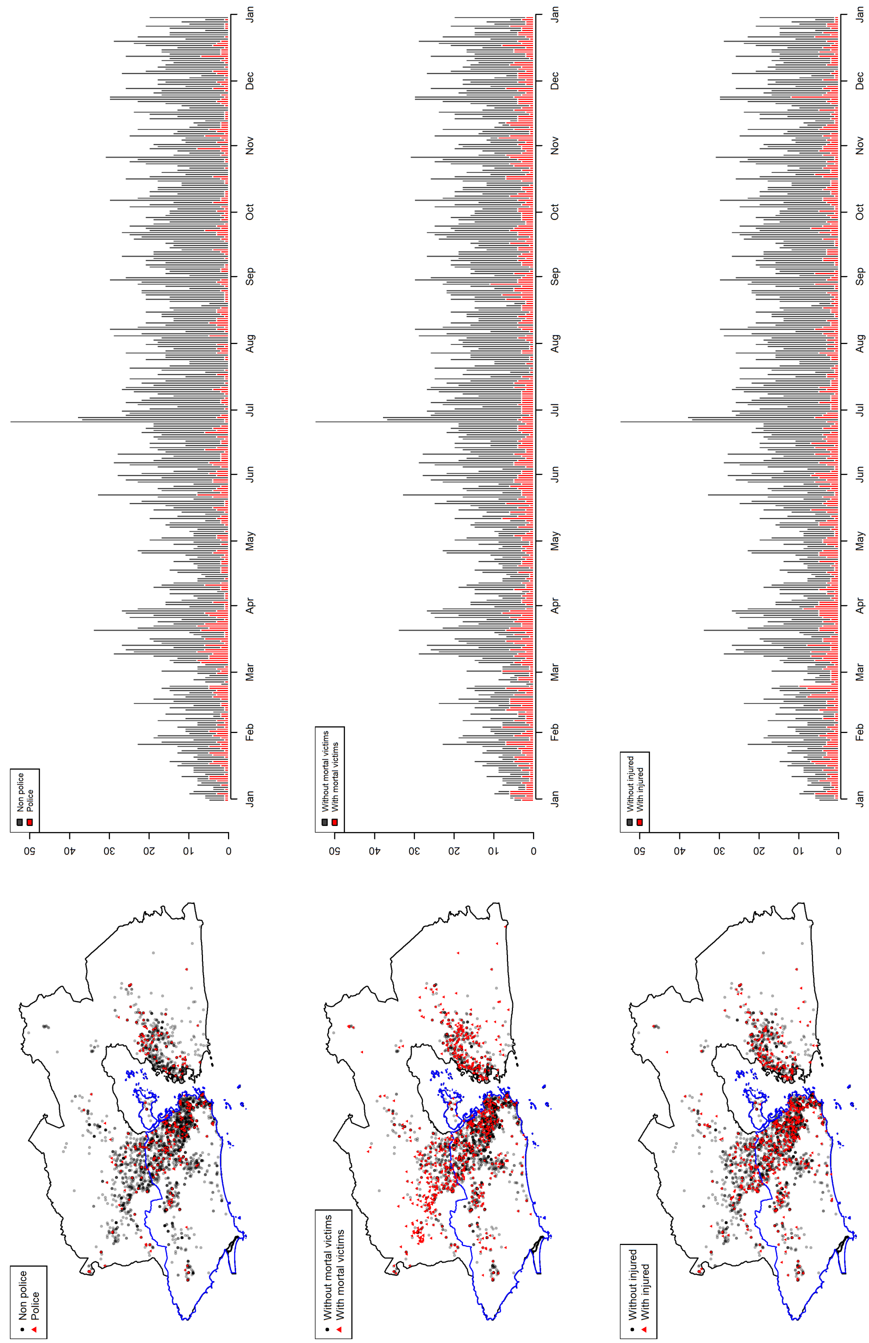

Figure 2: Spatial and temporal patterns of gunfire in Rio de Janeiro (region in blue) and its metropolitan area during 2017. Gun shootings with police interventions (top), mortal victims (center) and injured (bottom) in red.

area. Gunfire violence affected 165804 students in Rio´s municipal school system, which suffered at least one day of closure. Acari, with 45 days of suspension out of a total of 198 in the academic year 2017, was the most affected area.

3 Nonparametric point processes analysis

Point processes are mathematical models that govern the occurrence of a random number of events on a bounded domain, $W \subset \mathbb{R}^d$. 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 point processes according with the type of event. Spatial point processes generate a random number of events $\mathbf{X} = \{\mathbf{x_1}, \ldots, \mathbf{x_N}\}$ in a planar region $W \subset \mathbb{R}^2$ with area $|W| > 0$. Spatiotemporal point processes comprise the location and time of occurrence of a random number of events, $\mathbf{S} = \{(\mathbf{x_1}, \mathbf{t_1}), \ldots, (\mathbf{x_N}, \mathbf{t_N})\}$, irregularly placed in $W \times T \subset \mathbb{R}^2 \times \mathbb{R}^+$ Throughout this paper point processes and patterns are denoted in bold capitals and events are denoted in bold.

The *Fogo Cruzado* dataset introduced in Section 2.2, which reports the location and time of occurrence of gunfire in the Rio de Janeiro metropolitan area, can be seen as a realization of a spatiotemporal point process. Each event is marked by the number of mortal and injured police and civil victims. For simplicity, in this work we classify gun shootings into those with and without mortal and injured victims but do not account for the amount. Shootings are also divided into those with and without police intervention.

Nonparametric inference techniques recently developed for first and second-order analysis of spatial and spatiotemporal point processes allow us to answer important questions regarding the distribution and behavior of gunfire in the Rio de Janeiro metropolitan area. Kernel estimators of the first-order intensity functions (Diggle, 1985; Fuentes-Santos et al., 2016) characterize the spatial distribution of the different types of gunfire events under study, and the nonparametric comparison of first-order intensities (Fuentes-Santos et al., 2017) tests, for instance, whether gun shootings with and without mortal victims have the same spatial distribution or there is any area with higher mortality risk. A first-order spatiotemporal separability test (Fuentes-Santos et al., 2018) checks whether the spatial distribution of gunfire varies over time. Second-order characteristics, which analyze pairwise interactions between events, allow us to test if the gunfire patterns are clustered, regular or random, i.e., whether a shooting at a given location enhances (clustered) or inhibits (regular) gunfire occurrence in its neighborhood. Second-order characteristics of multitype point processes allow us to analyze the relationship between different types of events. We can test, for instance, if gun shootings with and without police interventions appear closer or further apart than expected if they were independent.

3.1 First-order analysis

First-order intensity estimation is a main issue in the analysis of any observed spatial point pattern. In our case, the first-order intensity describes the spatial distribution of gunfire in the Rio de Janeiro metropolitan area, and allows us to identify areas with high violence incidence, i.e., chronic hotspots. The first-order intensity function, $\lambda(x)$, is defined as

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

where $\mathbb{E}$ denotes expectation in both the number and location of events, $|dx|$ and $N(dx)$ denote the area and number of events of $\mathbf{X}$ in dx, which is an infinitesimal disc centered at location x. Intuitively. $\lambda(x)|dx|$ is the probability for dx to contain exactly one event of $\mathbf{X}$. A point process is homogeneous if its first-order intensity is constant, $\lambda(x) = \lambda > 0$, and inhomogeneous otherwise. The density of event locations is defined as $\lambda_0(x) = \lambda(x)/\int_W \lambda(x)dx$.

Diggle (1985) introduced the kernel intensity estimator for point processes in $\mathbb{R}$, whose extension to the spatial framework is natural and has been extensively used. The first-order intensity of gunfire patterns in this work is estimated through kernel smoothing with full bandwidth matrix (Fuentes-Santos et al., 2016)

$$\hat{\lambda}_H(x) = p_H(x)^{-1}|H|^{-1/2}\sum_{i=1}^{N} k\left(H^{-1/2}\left(x - \mathbf{x_i}\right)\right) \tag{2}$$

where $k(\cdot)$ is a Gaussian kernel, H, is a positively defined bandwidth matrix and $|H|$ denotes the determinant of H. We apply the 2-stages plug-in algorithm introduced by Fuentes-Santos et al. (2016) to select the optimal bandwidth matrix. Here, $p_H(x) = \int_W |H|^{-1/2}k\left(H^{-1/2}(x-y)\right)dy$ is the edge-correction term that reduces the bias near the boundary of the observation domain, W, and guarantee the consistency of the kernel density of event locations

$$\hat{\lambda}_{0,H}(x) = N^{-1}\hat{\lambda}_H(x)\, I\,(N > 0) \tag{3}$$

where $I(\cdot)$ denotes the indicator function (See Appendix A1 for further details).

The nonparametric test proposed by Fuentes-Santos et al. (2017) was used to compare the first-order intensity of different types of gunfire events. This procedure cna be used to test whether gun shootings with and without victims have the same spatial distribution. The first-order intensity of two spatial point processes, $\mathbf{X_1}$ and $\mathbf{X_2}$, with the same spatial distribution are proportional and consequently, their densities of event locations are equal. 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 these properties, Fuentes-Santos et al. (2017) extended the nonparametric test proposed by Duong et al. (2012) for multivariate data to the spatial point process framework, and proposed using a classical squared discrepancy measure to compare the spatial distribution of two observed spatial point patterns

$$T = \int_W \left(\lambda_{01}\left(x\right) - \lambda_{02}(x)\right)^2 dx = \psi_1 + \psi_2 - \left(\psi_{12} + \psi_{21}\right) \tag{4}$$

where $\lambda_{01}\left(x\right)$ and $\lambda_{02}\left(x\right)$ are the densities of event location of each spatial point process, $\psi_j = \mathbb{E}_{\mathbf{X_j}}\left[\lambda_{0j}\left(x\right)\right]$ and $\psi_{ij} = \mathbb{E}_{\mathbf{X_i}}\left[\lambda_{0j}\left(x\right)\right]$, for $i, j = 1, 2$. Using kernel smoothing with plug-in bandwidth to estimate each component of T (Chacón and Duong, 2010), we obtain the test statistic.

$$\hat{T} = \hat{\psi}_1 + \hat{\psi}_2 - \left(\hat{\psi}_{12} + \hat{\psi}_{21}\right) \tag{5}$$

The null distribution of $\hat{T}$ is asymptotically normal under regularity conditions analogous to those assumed in the classical multivariate distribution framework. However, the slow convergence rate to the normal distribution discourages using the asymptotic null distribution as calibration procedure. For this reason, Fuentes-Santos et al. (2017) proposed a smooth bootstrap algorithm, which computes the relative position of the test statistic of the observed pattern, $\hat{T}_1$, with respect to $B-1$ test statistics, $\{\hat{T}_b\}_{b=2}^{B}$ measuring the discrepancy between pairs of inhomogeneous Poisson point processes with expected number of events equal to those of the observed patterns and first-order intensity proportional to that of the unmarked point process $\mathbf{X} = \mathbf{X_1} \cup \mathbf{X_2}$. The empirical p-value of the test is the proportion of bootstrap T-statistics larger than $\hat{T}_1$. (See Appendix A2 for further details)

If the T-test finds differences between, for instance, gunfire with and without mortal victims, we may be interested on where did those differences occur. This, we need a local test that identifies areas with high and low mortality risk. This question can be addressed through the relative risk function, which was introduced by Bithell (1990) to compare the geographical distribution of disease cases, $\mathbf{X}_1$, and a random sample of the population at risk (controls), $\mathbf{X}_2$, and is defined as the ratio between their densities of event locations, $r(x) = \lambda_{01}(x)/\lambda_{02}(x)$. Given that $\lambda_{0j}(x), j = 1, 2$ are strictly positive, it is preferable working with the log-relative risk $\rho(x) = \log\left(\lambda_{01}(x)/\lambda_{02}(x)\right)$ to handle the asymmetry between the number of cases and controls. We have estimated the log-relative risk of police presence and victims using the symmetric adaptive kernel method proposed by Davies et al. (2016). The kernel log-relative risk is defined as $\hat{\rho}(x) = \log\left(\hat{\lambda}_{01,h}(x)/\hat{\lambda}_{02,h}(x)\right)$, where for $j = 1, 2$

$$\hat{\lambda}_{0j,h}\left(x\right) = \frac{1}{N_j}\sum_{i=1}^{N_j}\frac{1}{h_{ji}^2}k\left(\frac{x-\mathbf{x_{ji}}}{h_{ji}}\right)I\left(N_j > 0\right) \tag{6}$$

where $h_{ji} = h(\mathbf{x_{ji}})$ is a local bandwidth, which adapts the degree of smoothing to the number of events in the neigborghood of $\mathbf{x_{ji}}$ and, consequenlty, proivdes more accurate estimates of local features than a fixed bandwidth parameter. Following the proposal of Davies et al. (2016), we use a common bandwidth function $h(\mathbf{x}) = h_0\alpha(\mathbf{x})$ based on the unmarked point processes $\mathbf{X} = \mathbf{X_1} \cup \mathbf{X_2}$ to estimate the density of event locations of cases and controls. Thus $h\left(\mathbf{x}\right) = h_0\hat{\alpha}\left(\mathbf{x}\right)$, where the global bandwidth h_0 is obtained by optimal smoothing (Terrell, 1990), and following Abramson (1982) $\hat{\alpha}\left(\mathbf{x}\right) = \hat{\gamma}^{-1}\hat{\lambda}_{0,h_p}(x)^{-1/2}$ with pilot scalar bandwidth $h_p = 0.5h_0$, and $\hat{\gamma} = \exp\{N^{-1}\sum_{i=1}^{N}\hat{\lambda}_{0,h_p}\left(\mathbf{x_i}\right)\}$. Comparing expressions (A-3) and (A-8) we observe that the latter does not correct for edge effects. The edge corrector depends on the spatial locations and the bandwidth parameter,then, ths term cancels out in $\hat{\rho}(x)$ if the same bandwidth function is used to obtain the kernel density of event locations for for cases and controls. Finally, as proposed by Kelsall and Diggle (1995), we compute Monte-Carlo tolerance contours to identify areas with high mortality risk or police presence. To do so (i) we simulate $B-1$ pairs of inhomogeneous spatial point patterns with expected number of events equal to those of the observed case and control patterns and first-order intensity proportional to that of the unmarked point processes $\mathbf{X} = \mathbf{X_1} \cup \mathbf{X_2}$. (ii) we estimate the log-relative risk function for each pair of simulated case and control patterns $\{\hat{\rho}_b(x)\}_{b=2}^{B}$, (iii) given a significance level α we compute the tolerance surface, $\hat{\rho}_{1-\alpha}(x)$, as the $100\,(1-\alpha)$ percentile of the Monte-Carlo log-relative risk functions at each location, (iv) we have high risk at any location x where $\hat{\rho}\left(x\right) > \hat{\rho}_{1-\alpha}(x)$ (See Appendix A3 for further details).

Up to now we have focused on the spatial distribution of gunfire in the Rio de Janeiro metropolitan area, overlooking the temporal dimension. This gap can be addressed through first-order analysis of spatiotemporal point processes, which allows us to characterize the distribution of gunfire events in both space and time. The spatiotemporal intensity function is defined as a natural extension of the first-order intensity function of a spatial point process (Diggle, 2013)

$$\lambda(x,t) = \lim_{|dx \times dt| \to 0} \left\{ \frac{\mathbb{E}\left[N(dx,dt)\right]}{|dx \times dt|} \right\} \tag{7}$$

where $N(dx, dt)$ is the number of events in the volume $dx \times dt$, dx is an infinitesimal disc containing location x, and dt is an infinitesimal interval around time t. As well as in the spatial domain, we have applied kernel smoothing to estimate the spatiotemporal intensity of gunfire

$$\hat{\lambda}_{H_s,h_t}(x,t) = p_{H_s,h_t}(x,t)^{-1}|H_s|^{-1/2}h_t^{-1} \sum_{i=1}^{N} k_s\left(H_s^{-1/2}\left(x - \mathbf{x_i}\right)\right) k_t\left(h_t^{-1}\left(t - \mathbf{t_i}\right)\right) \tag{8}$$

where $k_s(\cdot)$ and $k_t(\cdot)$ are bivariate and univariate Gaussian kernels, H_s is the two-dimensional bandwidth matrix for the spatial component, and h_t is the bandwidth for the temporal component. $p_{H_s,h_t}(x,t) = \int_T \int_W k_{s,H_s}(x-u)k_{t,h_t}(t-v)dudv = \int_W k_{s,H_s}(x-u)du \int_T k_{t,h_t}(t-v)dv$ is the spatio-temporal edge-corrector, where $p_{H_s}(x) = \int_W k_{s,H_s}(x-u)du$ and $p_{h_t}(t) = \int_T k_{t,h_t}(t-v)dv$ represent, respectively, the bivariate edge-corrector for spatial locations and the univariate edge-corrector for the temporal component.

Testing whether gunfire patterns are separable, i.e., whether their spatiotemporal intensity function can be decomposed into the product of its spatial and temporal marginals,

$\lambda(x;t) = \lambda_1(x)\lambda_2(t)$, is crucial for two reasons. On one hand, the separabilty test allows us to determine whether the spatial distribution of gunfire remained constant during 2017 or varied over time. On the other hand, separability simplifies its estimation avoiding the curse of dimensionality. Taking into account that in a separable point process the risk of observing an event at time t is spatially invariant, i.e., the ratio between the intensity functions of the spatiotemporal point process and its spatial marginal does not depend on event locations, Fuentes-Santos et al. (2018) proposed testing separability through a regression test that checks if the log-ratio function $\rho(x,t) = \log\left(\lambda(x,t)/\lambda_1(x)\right)$ depends on the spatial locations. To implement the test we first estimate $\rho(x,t)$ as the log-ratio of the kernel spatiotemporal (8) and spatial (2) intensity functions with diagonal bandwidth matrices selected by least-squares cross-validation, as the lack of robust estimates for the reciprocals of density functions hampers the use of a plug-in bandidth selctor. (see details in Fuentes-Santos et al. (2018)). Once obtained $\hat{\rho}(x,t)$, we consider a regression problem where the log-ratio function evaluated at each event is a response variable, $Y = \{y_i = \hat{\rho}(\mathbf{x_i}, \mathbf{t_i}), i = 1,\ldots,n\}$, that may depend on the spatial covariate $X = \{x_i = (x_{i1}, x_{i2}), i = 1,\ldots,n\}$ comprising the event locations, and we test for the effect of X on Y. Following Bowman and Azzalini (1997), we shall discriminate between two competing models

$$\mathcal{H}_0 : E\left[Y_i|X_i\right] = \mu. \quad \mathcal{H}_1 : E\left[Y_i|X_i\right] = m(X_i)$$

We estimate μ by the empirical mean, $\hat{y} = n^{-1}\sum_{i=1}^{n} Y_i$, and the unknown smooth function, $m(x)$, by kernel regression. We compute the residual sum of squares for the null, RSS_0, and alternative, RSS_1, models and define the generalized test statistic, F, as follows

$$F = \frac{(RSS_0 - RSS_1) / (df_1 - df_0)}{RSS_1 / df_1} \quad (9)$$

where df_0, df_1 denote the degrees of freedom for these residuals. The test is calibrated through a permutation test, which relies on the fact that under $\mathcal{H}_0$ the pairing of any particular x and y is completely random. Then, the distribution of the test statistic, F, can be generated by simulation, computing the test statistic for $B - 1$ random pairings of the observed values of X and Y. The empirical p-value of the test is the proportion of simulated F-statistics larger than that obtained for the observed data (See Appendix A4 for further details).

3.2 Second-order analysis

Once characterized the spatial and spatiotemporal distribution of gunfire patterns, we focus on searching for interactions between events. The reduced second moment measure or K-function (Ripley, 1977) has been commonly used to describe the dependence structure of spatial point patterns, i.e to check whether events are independent (Poisson point process) or there is any kind of interaction between them. The K-function of a homogeneous spatial point process is defined as

$$K(r) = \lambda^{-1} \mathbb{E}[N_0(r)] \quad (10)$$

where $N_0(r)$ is the number of further events within distance r of an arbitrary event.

The homogeneity assumption can be very restrictive and unrealistic when dealing with real data. For this reason Baddeley et al. (2000) extended the K-function to the case of second-order reweighted stationary (SORS) point processes, which first-order intensity is inhomogeneous and bounded away from 0 and second-order properties depend on the distance between events but not on their locations, and defined the inhomogeneous K-function as follows

$$K_{inhom}(r) = \frac{1}{|W|}\mathbb{E}\left[\sum_{\mathbf{x_i},\mathbf{x_j}\in\mathbf{X}\bigcap W}^{i\neq j} \frac{I\left(\parallel \mathbf{x_i}-\mathbf{x_j} \parallel < r\right)}{\lambda(\mathbf{x_i})\lambda(\mathbf{x_j})}\right] \tag{11}$$

which natural empirical estimator is

$$\hat{K}_{inhom}(r) = \frac{1}{|W|}\sum_{\mathbf{x_i},\mathbf{x_j}\in\mathbf{X}\bigcap W}^{i\neq j} \frac{I\left(\parallel \mathbf{x_i}-\mathbf{x_j} \parallel \leq r\right)}{\hat{\lambda}\left(\mathbf{x_i}\right)\hat{\lambda}\left(\mathbf{x_j}\right)}\omega_{i,j}^{-1};\ 0 \leq r \leq r_{max} \tag{12}$$

where $I(\cdot)$ is the indicator function, r_{max} is the maximum distance at which the function is evaluated. w_{ij}, defined as the reciprocal of the proportion of the disc centered at location $\mathbf{x_i}$ with radius $\parallel \mathbf{x_i}-\mathbf{x_j} \parallel$ that lays inside W, is Ripley´s edge-corrector. $\hat{\lambda}(x)$ is an estimator of the first-order intensity, in this work the kernel estimator with plug-in bandwidth matrix defined in expression (2). The L-function (Besag, 1977), defined as $L_{imhom}(r) = \sqrt{K_{inhom}(r)/\pi}$, is a transformation of the K-function widely used in practice.

The K-function of a spatial Poisson point processes is $K(r) = \pi r^2$, consequently $L(r) = r$, whereas $K(r) < \pi r^2$ and $K(r) > \pi r^2$ indicate inhibition or clustering between events at distance r, respectively. This property suggests using the L-function as an statistic in a Monte-Carlo test to check whether gunfire events occurred independent or they had a clustered or regular distribution. The Monte-Carlo test is implemented as follows: (i) compute the empirical L-function, $\hat{L}_{inhom,1}(r)$, for the observed pattern; (ii) obtain $\{\hat{L}_{inhom,b}(r)\}_{b=2}^{B}$ for $B-1$ realizations of the inhomogeneous spatial Poisson point process with the same first-order intensity as the observed pattern, (iii) compute the upper and lower envelopes of the simulated L-functions

$$\hat{L}_{inhom}^{lo} = \min_{b=2,\ldots,B} \hat{L}_{inhom,b}(r) \qquad \hat{L}_{inhom}^{hi} = \max_{b=2,\ldots,B} \hat{L}_{inhom,b}(r)$$

and (iv) gunfire events are independent if $\hat{L}_{inhom,1}(r)$ falls within the upper and lower envelopes, clustered or regular when $\hat{L}_{inhom,1}(r)$ is above the upper or below the lower envelope, respectively. The maximum distance, r, at which $\hat{L}_{inhom,1}(r)$ is outside the envelopes is the interaction radius. As usual in practice, we consider $B = 40$ to get a significance level $\alpha = 0.05$ when the envelopes are given by the minimum and maximum values of the statistic in the Monte-Carlo simulations.

We may also wonder whether different types of gunfires, e.g. those with and without mortal victims, are independent or occur closer or further apart than expected if they were independent. This question can be answered by Monte-Carlo tests based on the K-cross function (Ripley, 1981), which is a natural extension of the K-function to the multitype framework. The K-cross of a homogeneous spatial point process is

$$K_{ij}(r) = \lambda_j^{-1}\mathbb{E}\left[N_{0ij}(r)\right] \tag{13}$$

where $N_{0ij}(r)$ is the number of type j events within distance r of an arbitrary type i event. If any of the marginal point processes is inhomogeneous we define the inhomogeneous K-cross as an extension of the inhomogeneous K-function (11), and its empirical estimated as follows

$$\hat{K}_{inhom,ij}(r) = \frac{1}{|W|} \sum_{\substack{\mathbf{x_{i,k}} \in \mathbf{X_i} \bigcap W \\ \mathbf{x_{j,l}} \in \mathbf{X_j} \bigcap W}} \frac{I\left(\| \mathbf{x_{i,k}} - \mathbf{x_{j,l}} \| < r\right)}{\hat{\lambda}_i(\mathbf{x_{i,k}})\hat{\lambda}_j(\mathbf{x_{j,l}})} \omega\left(x_{i,k}, x_{j,l}\right)^{-1} \;\; 0 \leq r \leq r_{max} \qquad (14)$$

where the edge-corrector $\omega\left(x_{i,k}, x_{j,l}\right)$ is analogous to ω_{ij} in expression (12). $\hat{\lambda}_i(x)$ and $\hat{\lambda}_j(x)$ are the kernel estimators of the first-order intensity functions for type i and type j spatial point patterns. As well as in the univariate case, $K_{inhom,ij}(r) = \pi r^2$ if $\mathbf{X}_i$ and $\mathbf{X}_j$ are independent, while values above or below this threshold indicate attraction or inhibition between type i and type j events, respectively. Therefore, we can test for dependence between two types of gunfire events through a Monte-Carlo test analogous to the L-test outlined above.

The Monte Carlo tests introduced above analyze the spatial interaction between gun shootings without accounting for the temporal lag between them. However, testing for spatiotemporal interaction, i.e., testing if a gun shooting increases or reduces the risk of new events on its neighborhood during the next hours or days, is crucial for a proper analysis of crime behavior and to detect near-repeat effects. To address this issue we have applied a Monte-Carlo test based on the spatiotemporal inhomogeneous K-function (Gabriel and Diggle, 2009), the natural extension of (9) to the spatiotemporal framework, which is defined as follows

$$K_{ST,inhom}(r,t) = \frac{1}{|W \times T|}\mathbb{E}\left[\sum_{\substack{\mathbf{x_i},\mathbf{x_j}\in W\\ \mathbf{t_i},\mathbf{t_j}\in T}}^{i\neq j} \frac{I\left(\| \mathbf{x_i}-\mathbf{x_j} \| < r, |\mathbf{t_i}-\mathbf{t_j}| < t\right)}{\lambda((\mathbf{x_i},\mathbf{t_i}))\lambda((\mathbf{x_j},\mathbf{t_j}))}\right] \tag{15}$$

$K_{ST,inhom}(r,t) = \pi r^2 t$ for a Poisson point process, while larger and smaller values indicate, respectively, clustering and inhibition (Gabriel and Diggle, 2009). To implement the spatiotemporal K-test we first obtain the empirical estimator of $K_{ST,inhom}(r,t)$

$$\hat{K}_{ST,inhom}(r,t) = \frac{1}{|W \times T|}\sum_{\substack{\mathbf{x_i},\mathbf{x_j}\in W\\ \mathbf{t_i},\mathbf{t_j}\in T}}^{i\neq j} \frac{I\left(\| \mathbf{x_i}-\mathbf{x_j} \| < r, |\mathbf{t_i}-\mathbf{t_j}| < t\right)}{\hat{\lambda}\left(\mathbf{x_i},\mathbf{t_i}\right)\hat{\lambda}\left(\mathbf{x_j},\mathbf{t_j}\right)}\omega_{i,j}^{-1} \tag{16}$$

$$0 \leq r \leq r_{max} \; 0 \leq t \leq t_{max}$$

where $I(\cdot)$ is the indicator function, ω_{ij} is Ripley´s spatiotemporal edge-corrector, r_{max}, t_{max} are the maximum spatial and temporal distances at which the function is evaluated, and $\hat{\lambda}(x,t)$ is an estimator of the spatiotemporal first-order intensity function. Gabriel and Diggle (2009) assume first-order separability, i.e., they use a separable estimator of the spatiotemporal intensity function, $\hat{\lambda}(x,t) = \hat{\lambda}_1(x)\hat{\lambda}_2(t)$, and attribute any non separable effect to the second-order structure. We have used separable or nonseparable kernel intensity estimated in agreement with the results of the first-order separability test introduced in Section 3.1. As well as in the spatial framework, the Monte-Carlo test compares the spatiotemporal K-function of the observed pattern, $\hat{K}_{ST,inhom}(x,t)$, with upper and lower envelopes determined by Monte-Carlo realizations of spatiotemporal Poisson processes with the same intensity as the observed pattern. The maximum values r and t for which $\hat{K}_{ST,inhom}(x,t)$ falls above or bellow the envelopes determine respectively the spatial and temporal interaction radius.

3.3 Software

The statistical analysis of the gunfire spatial and spatiotemporal patterns was conducted with the R statistical software (R Core Team, 2018). The *Spatstat* package (Baddeley et al., 2015) was used for kernel intensity estimation and second-order analysis of spatial point patterns, whereas plug-in bandwidths were obtained with the help of the *ks* package (Duong, 2018). The *kde.test* function in the *ks* package was adapted by (Fuentes-Santos et al., 2017) to implement the T-test for comparison of first-order intensities. Kernel log-relative risk functions and their tolerance contours were estimated using the *sparr* package (Davies et al., 2018). Fuentes-Santos et al. (2018) extended some functions in the *sparr* package (Davies et al., 2018) to estimate the spatiotemporal intensity and log-ratio functions to implement the separability test with the help of the *sm* package (Bowman and Azzalini, 2014).

4 Results

In this section we outline the results of the first and second-order analysis of gun shootings registered in the Rio de Janeiro metropolitan area during 2017.

4.1 Spatial and spatiotemporal distribution of gunfire

Figure 3 shows the kernel intensity (left) and kernel density (right) estimators for the spatial and temporal patterns of gunfire reports in the Rio de Janeiro metropolitan area. We

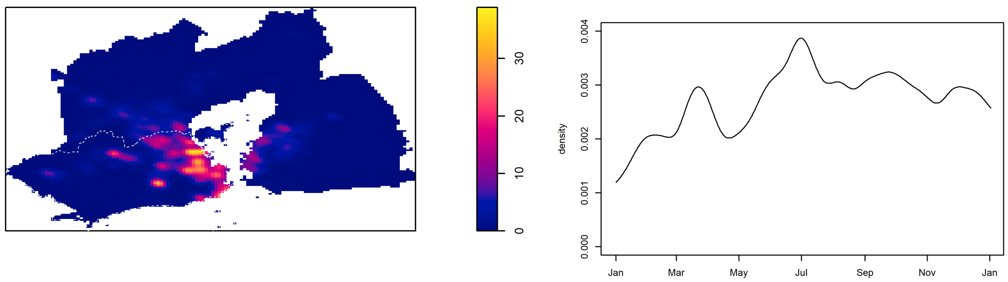

Figure 3: Kernel estimator for the first-order intensity and edge-corrected kernel estimatior of the temporal density of gunfire in the Rio de Janeiro metropolitan area.

observe a clear inhomogeneous pattern with hotspots in Niteroi and São Gonçalo, although the highest gunfire incidence was registered in the eastern area of Rio de Janeiro, which includes the favelas complexes in the *Zona Norte*. The temporal pattern of gunfire is also inhomogeneous, Figure 3 (right) shows higher gunfire incidence in the second semester of 2017, after the peak registered in July. In order to test whether the spatial distribution of gunfire varied over time we have compared the first-order intensity of gun shootings by months (Figure 4, Table 2), the T-test found differences between months except for January and February, which had similar spatial distribution. Finally, the nonparametric spatatiotemporal separability test confirms that the spatial distribution of gunfire violence in the Rio de Janeiro metropolitan area varied over time (F-test, p-value < 0.005), see the spatiotemporal intensity in B.1 (Appendix B).

Comparison between the first-order intensity of gunfire events with and without police intervention (Figure 5) can be used, for instance, by police departments to analyze their patrolling strategies. The T-test (p-value < 0.005) found significant differences be-

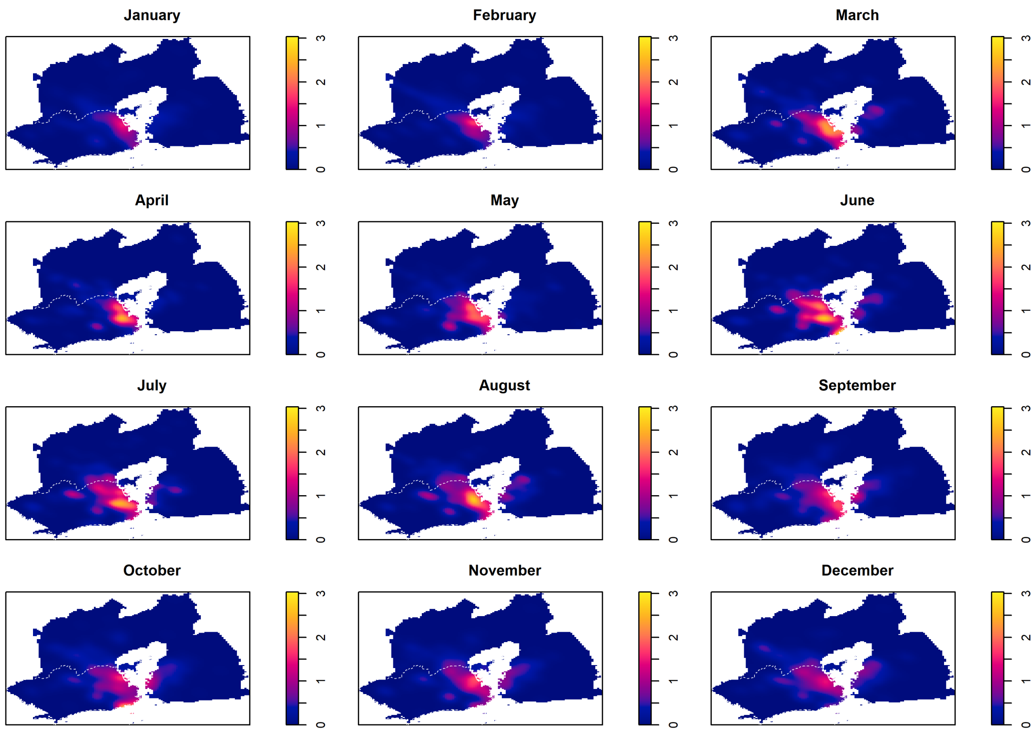

Figure 4: Kernel intensity function of gunfire in the Rio de Janeiro metropolitan area by month.

Table 2: P-values for pairwise nonparametric comparison of gunfire intensities between months. T-test (A-7) with $B = 200$ realization of the null hypothesis for bootstrap calibration

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov
Feb	0.399										
Mar	0.041	<0.005									
Apr	<0.005	<0.005	<0.005								
May	<0.005	<0.005	<0.005	<0.005							
Jun	<0.005	<0.005	<0.005	<0.005	<0.005						
Jul	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005					
Aug	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005				
Sep	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005			
Oct	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005		
Nov	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	
Dec	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005

tween their spatial distributions. The log-relative risk function comparing the intensity of gunfire with (cases) and without (controls) police intervention (Figure 5, bottom left) shows high police presence in some hotspots but also in areas with low gunfire incidence (Figure 3). We also observe differences between the temporal patterns (p-value < 0.005), with a relative decrease in police interventions in the second semester of 2017. Note that this period suffered higher gunfire activity than the first semester (Figure 3). Finally, the F-test found that both gunfire patterns are nonseparable (p-value < 0.005), see the spatiotemporal intensities of gunfire with and without police intervention in B.2 (Appendix B).

Now we focus on the severity of gunfire violence through the analysis of gun shootings with mortal (Figure 6) and injured (Figure 7) victims. Gunfire with and without mortal victims have different spatial distribution (T-test, p-value < 0.005, Figure 6 top). In par-

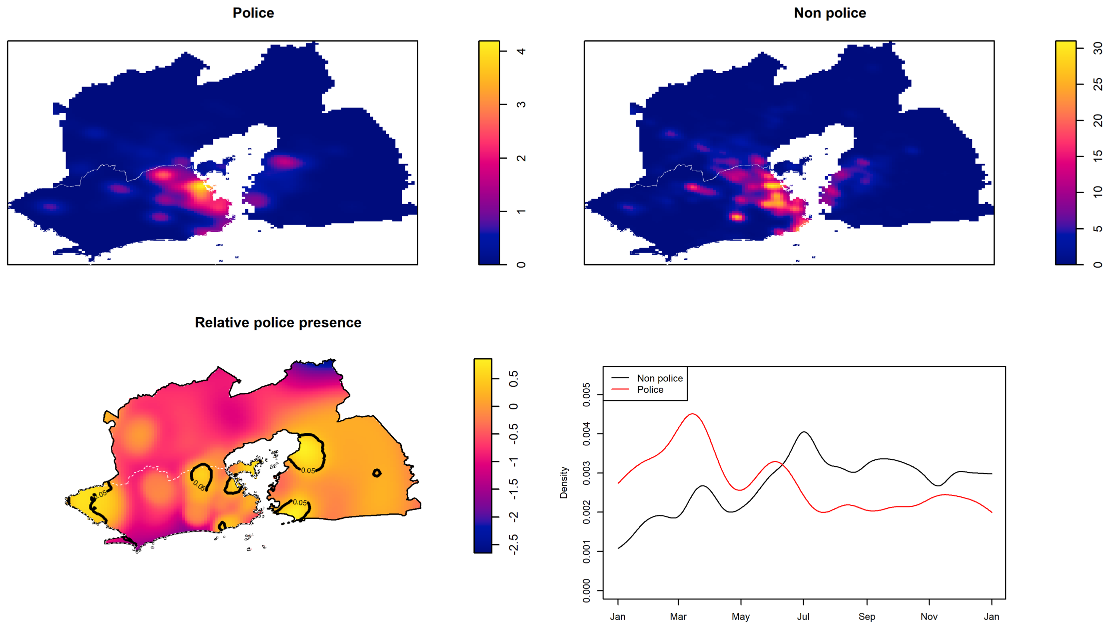

Figure 5: Top: kernel intensity function for the spatial patterns of gun shootings with and without police intervention (Figures have different scale). Bottom left: kernel log-relative risk function with Monte-Carlo tolerance contours ($B = 200$, $\alpha = 0.05$) for high police presence (black lines). Bottom right: kernel density estimator for the temporal pattern of gunfire events with and without police intervention.

ticular, the log-relative risk function (Figure 6, bottom left) reports high mortality in areas with low gunfire activity, and low risk in the eastern area of Rio de Janeiro, which suffered extremely large gun violence during 2017 (Figure 3). We also observe differences between the temporal patterns (p-value < 0.005), with a relative mortality decrease between April and August followed by an increase from September onward. The F-test rejected separability for gunfire with and without mortal victims (p-value < 0.005), i.e., their spatial distribution varied over time, as can be seen in B.3 (Appendix B).

The T-test detected significant differences between the spatial distribution of gun shootings with and without injured victims (p -value < 0.005). As well as for fatalities, Figure 7 (bottom, left) shows high risk of injuries in areas with low gunfire incidence. We also found differences between the temporal patterns ($p - value < 0.005$), with higher density of gun shootings with injured victims during the first semester of 2017 and low relative incidence thereafter. The F-test found that gunfire with and without injuries have nonseparable first-order structure ($p - value < 0.005$, see B.4 in Appendix B).

4.2 Interaction between gun shootings

Once characterized the spatial and spatiotemporal distribution of gunfire violence in the Rio de Janeiro metropolitan area, our next aim is searching for spatial and spatiotemporal interactions between gun shootings. For this purpose, we have applied the Monte-Carlo tests outlined in Section 3.2. We have tested for spatial and temporal interactions up to $r_{max} = 5$ km and $t_{max} = 15$ days, respectively.

The inhomogeneous L-test (Figure 8, left) detected spatial clustering between gun shootings at small distances ($r < 0.5$ km). The spatiotemporal K-test (Figure 8, right) indicates that this aggregation occurs within 1 day, i.e., the occurrence of a gunfire at a given location increases the risk of suffering more events within 0.5 km during the next 24 hours.

The inhomogeneous L-test also found attraction up to 1 km for gunfire with and without police intervention (Figure 9, top). Similar results were found when testing for interactions

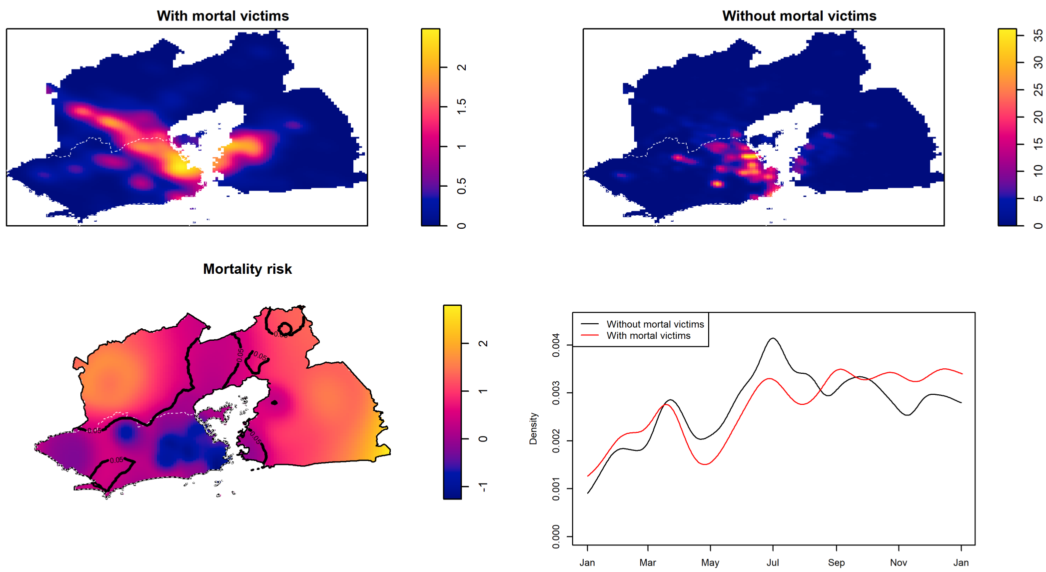

Figure 6: Top: kernel intensity function for the spatial patterns of gunfire events with and without mortal victims (Figures have different scale). Bottom left: Bottom left: kernel log-relative risk function with Monte-Carlo tolerance contours ($B = 200$, $\alpha = 0.05$) for high mortality risk (black lines). Bottom right: edge-corrected kernel density estimator for the temporal pattern of gunfire with and without mortal victims.

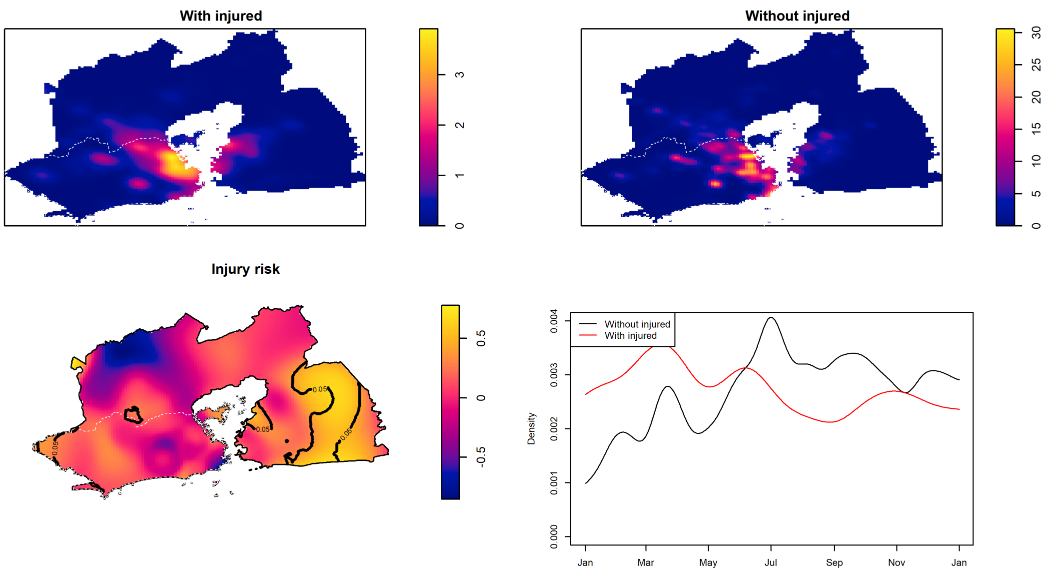

Figure 7: Top: kernel intensity function for the spatial patterns of gunfire events with and without injuries (Figures have different scale). Bottom left: relative risk function with Monte-Carlo tolerance contours ($\alpha = 0.05$) for high injury risk (black lines). Bottom right: edge-corrected kernel density estimator for the temporal pattern of gunfire with and without injured victims.

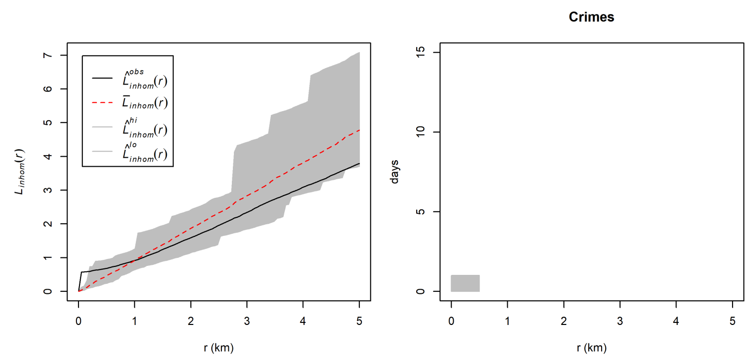

Figure 8: Left: inhomogeneous L-test for spatial interactions between gunfire events. Right: inhomogeneous K-test for spatiotemporal interactions between gunfire events, in grey spatiotemporal volume with $K_{ST,obs}(r,t) > K_{ST,hi}(r,t)$, where $K_{ST,hi}(r,t)$ is the upper envelope for $B - 1 = 39$ realizations of a spatiotemporal Poisson point process with the same first-order intensity as the observed pattern.

between gunfire events with and without police intervention (Figure 9, bottom left). The spatiotemporal K-test found long-term clustering in police interventions up to 0.5 km, which may indicate that there are some neighborhoods with constant police presence. The clustering radius increases to 1.5 km within 36 hours, this may indicate an intensification of police presence in a given neighbourhood after a gun shooting is reported in order to reduce near-repeat phenomena.

The second-order analysis of gunfire with and without mortal victims (Figure 10) detected spatial clustering in the marginal patterns, and between gun shootings with and without mortal victims at small distances ($r < 1$ km). The spatiotemporal K-test shows long-term clustering in gun shootings with mortal victims up to 0.5 km, this radius in-

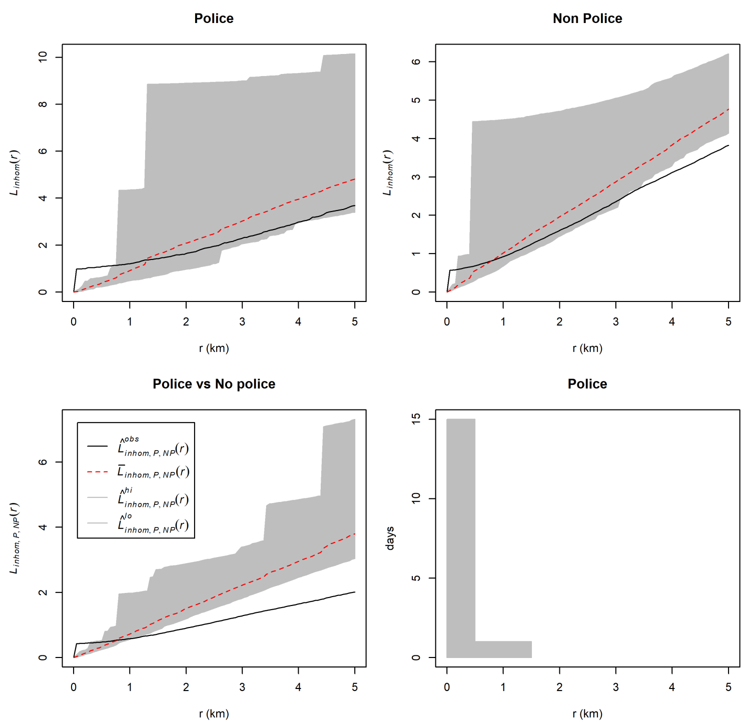

Figure 9: Inhomgoeneous L-test for gunfire with or without police intervention (top), L-cross test for interaction between crimes with and without police intervention (bottom left), and Inhomogeneous K-test for spatiotemporal interactions between crimes with police interventions (bottom right), see details in the caption of figure 8)

creases up to 1 km within 6 days after the occurrence of an event with mortal victims.

The inhomogeneous L-test (Figure 11) detected clustering in the spatial patterns of gun shootings with and without injuries, and positive interaction between them at small distances ($r < 1$ km). The spatiotemporal K-test found that the occurrence of a gunfire with injured victims increases the risk of new events with injuries within 0.5 km during the next 6 days.

5 Discussion

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 metropolitan area (Brazil), which has been suffering a continuous increase of violent crime 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.

This work analyzes the gunfire reports collected by *Fogo Cruzado* in 2017 that, in addition to the spatial location and time of occurrence, contain information about police presence and victims. Previous application of point process methods in criminology have

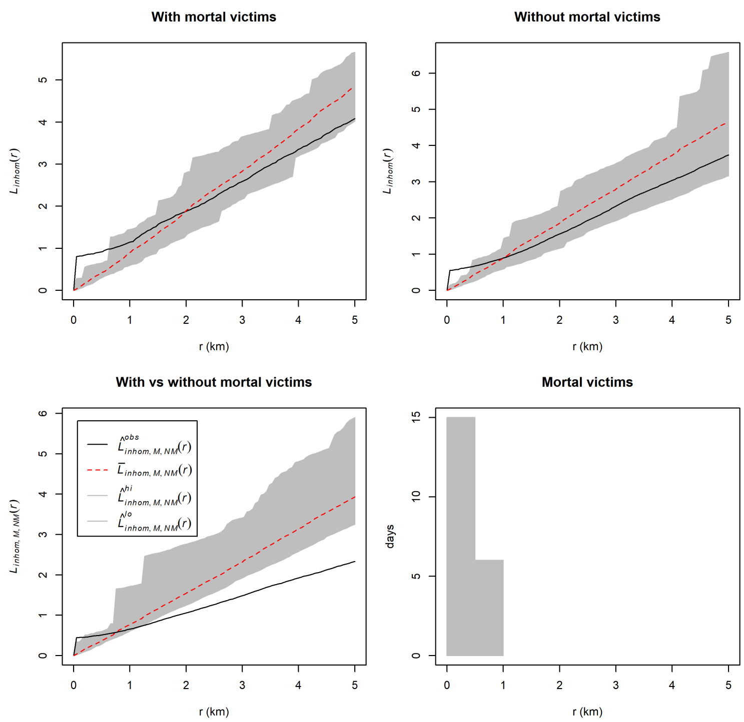

Figure 10: Inhomgoeneous L-test for gunfire with or without mortal victims (top), L-cross test for interaction between gunfire with and without mortal victioms (bottom left), and Inhomogeneous K-test for spatiotemporal interactions between gun shootings with morlal vicitoms (bottom right, see details in the caption of figure 8)

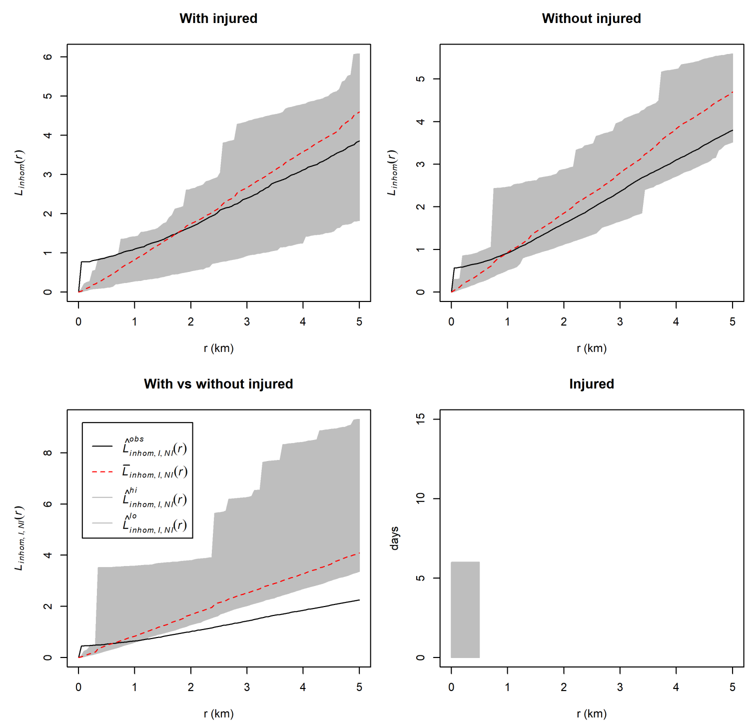

Figure 11: Inhomgoeneous L-test for gunfire with or without injured victims (top), L-cross test for interaction between gunfire with and without injured victioms (bottom left), and Inhomogeneous K-test for spatiotemporal interactions between gun shootings with injured vicitoms (bottom right, see details in the caption of figure 8)

mainly used kernel density estimation (Bowers et al., 2004; Chainey et al., 2008; Gerber, 2014) or self-exciting point process modeling (Mohler et al., 2011; Mohler, 2014; Reinhart and Greenhouse, 2018; Zhuang and Mateu, 2018) to predict future crime distribution. Using any of these methods implies assuming some hypothesis for the observed patterns, kernel density estimation assumes that events occur independently, whereas self-exciting point process modeling assumes a contagious like behavior. However, as this is the first spatiotemporal approach to *Fogo Cruzado* data, and we do not have enough information about gunfire dynamics in Rio de janeiro, we need a complete exploratory analysis prior to apply any of these models. For this reason we have applied nonparametric inference for spatial and spatiotemporal point processes to characterize the distribution of gunfire in the Rio de Janeiro metropolitan area, and test for interactions between gunfire events.

We have applied kernel intensity estimation to characterize the spatial distribution of gunfire. This method is analogous to the kernel hotpsot maps used by Chainey et al. (2008), but incorporates an edge-correction term, as the observation domain is bounded. The kernel intensity of gunfire identifies a hotspot in the eastern area of Rio de Janeiro, this area includes the *Zona Norte*, whose high violence rates may be linked to the presence of drug traffic gangs (Arias and Barnes, 2017). Although we also observe high police presence and hotspots for mortal and injured victims in the eastern area of Rio, the nonparametric comparison between point processes (Fuentes-Santos et al., 2017) found differences between the first-order structure of gunfire with police intervention or victims and those without them. In particular the relatively low mortality risk observed in the hotspot of the eastern area, which contrasts with the large risk in areas with low gunfire activity, suggests the need of a deeper analysis to find which factors increase mortality risk in the latter. Information about the cause of gun shootings may be useful for this purpose.

The spatiotemporal separability test (Fuentes-Santos et al., 2018) found that the spatial distribution of gun shootings, police intervention and victims varied over time. Therefore, in addition to static environmental or demographic factors, some time varying covariate, such as precursory crimes, need to be considered for an accurate modeling of gunfire spatiotemporal dynamics.

Taking into consideration the results of the separability test, we have applied an inhomogeneous spatiotemporal K-test with nonseparable kernel intensity, in contrast with the common practice which assumes first-order separability and attributes any nonseparable effect to the second-order structure (Gabriel and Diggle, 2009). The spatiotemporal K-test found clustering between gun shootings at short distances ($r = 0.5$ km) within 24 hours suggesting short-term contagious like effects. We have also found clustering for gun shootings with police interventions, mortal and injured victims. The temporal interaction radius for victims is larger than the observed for the whole gunfire pattern, suggesting that it is more difficult to control retaliatory effects as violence gets more severe. From the methodological point of view, these results suggest that a self-exciting point process model (Mohler et al., 2011) would be an accurate approach for the spatiotemporal behavior of gunfire in Rio de Janeiro.

In conclusion, this work provides a first analysis of violent crime data obtained from real time reports of users in a mobile app. Social media, in particular twitter has been used by Gerber (2014) as covariate in the analysis of crime data provided by Chicago police department. Nonparametric estimators and tests for first and second-order properties of point processes allowed us to describe the dynamics of gunfire in the Rio de Janeiro metropolitan area and provide valuable information for future research. The first-order analysis revealed that gun shootings have a highly inhomogeneous spatial distribution, the spatial distribution of police presence differs from that of gunfire, and the relative risk of mortal and injured victims is also inhomogeneous. We have also seen that the spatial distribution of gunfire, police intervention and victims varies over time. Second-order analysis suggests short-term near repeat effects in the gunfire patterns. Taking into account all this information, we can consider developing a self-exciting point process model with nonseparable background to forecast gunfire in the Rio de Janeiro metropolitan area. Visual inspection of the spatial first-order intensity may help criminologists to select environmental, demographic or socioeconomic covariates to include in the background component, as done by Reinhart and Greenhouse (2018), but we also need some short-term varying covariate to model the nonseparable gunfire dynamics. In view of the differences found between gunfire with and without police intervention and victims, we shall need different covariates to model each type of gunfire. Finally, the different spatial and temporal interaction radius found for each type of gunfire indicates differences in their near repeat behavior. Thus, in case of using leading indicators, such as precursory crimes, in the self-exciting component of the model, as done by Mohler (2014), we should select different indicators for each type of gunfire.

APPENDIX A. Nonparametric inference for spatial and spatiotemporal point processes

Point processes are mathematical models that govern the occurrence of a random number of events on a bounded domain, $W \subset \mathbb{R}^d,\ d \geq 1$. 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 point processes according with the type of event. Spatial point processes generate a random number of events $\mathbf{X} = \{\mathbf{x_1}, \ldots, \mathbf{x_N}\}$ in a planar region $W \subset \mathbb{R}^2$ with area $|W| > 0$. Spatiotemporal point processes comprise the location and time of occurrence of a random number of events, $\mathbf{S} = \{(\mathbf{x_1}, \mathbf{t_1}), \ldots, (\mathbf{x_N}, \mathbf{t_N})\}$, irregularly placed in $W \times T \subset \mathbb{R}^2 \times \mathbb{R}^+$.

This work analyses the spatiotemporal point pattern of gun shootings recorded by *Fogo Cruzado* in the Rio de Janeiro metropolitan area during 2017. In addition to the location and time of occurrence of each event, we also have information about police interventions, mortal victims and injuries. In this first approach to the *Fogo Cruzado* data set our aim is to characterize the distribution of gunfire and check if there are interactions between gun shootings in order to understand the dynamics of violent crime in Rio de Janeiro metropolitan areas. As pointed out by Diggle (2013), we cannot distinguish between hetereogeneity and interaction between events in an observed pattern when, as in this case, we lack of additional information such as covariates or a parametric model. Therefore, following the common practice in these situations, we first focus on first-order analysis assuming that the point process is Poisson, i.e., assuming that events are independent, and test for interactions between events once the first-order structure has been characterized. Bellow we provide detailed information about the nonparametric first and second-order inference methods applied in this work.

APPENDIX A.1. Kernel intensity estimation

Let $\mathbf{X}$ be a spatial point process defined in $\mathbb{R}^2$ and $\{\mathbf{x}_1, \ldots, \mathbf{x}_1\}$ a realization of $\mathbb{X}$ observed on a bounded region $W \subset \mathbb{R}^2$. The spatial distribution of events in the observation domain is characterized by the first-order intensity function, $\lambda(x)$, defined as follows

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

where $\mathbb{E}$ denotes expectation in both the number and location of events, $|dx|$ and $N(dx)$ denote the area and number of events of $\mathbf{X}$ in dx, which is an infinitesimal disc centered at location x. Intuitively. $\lambda(x)|dx|$ is the probability for dx to contain exactly one event of $\mathbf{X}$. A point process is homogeneous if its first-order intensity is constant, $\lambda(x) = \lambda > 0$, and inhomogeneous otherwise.

The first-order intensity can be estimated assuming a parametric model and estimating the unknown parameters by maximum pseudolikelihood (Waagepetersen, 2007). However, this procedure can lead to unreliable estimates if the assumed parametric model deviates from the true intensity function. A better alternative when, as in our case, we do not have enough information to define a parametric model is using a nonparametric estimator. Diggle (1985) introduced the kernel intensity estimator for one-dimensional point processes, which has been directly extended to the spatial domain as follows:

$$\hat{\lambda}_h(x) = \frac{1}{p_h(x)} \sum_{i=1}^{N} k_h\left(x - \mathbf{x_i}\right), \quad x \in \mathbb{R}^2 \tag{A-2}$$

where h is the bandwidth parameter, k denotes a bivariate kernel function, $k_h(x) = h^{-2}k\left(x/h\right)$ is the smoothed kernel, and $p_h(x) = \int_W h^{-2}k(h^{-2}(x-y))dy$ is the edge correction term. This term reduces the bias of the kernel intensity estimator near the boundary of the observation domain, which appears because events occurring outside W are not observed and, consequently not included in the estimator. The kernel intensity estimator has been limited to exploratory analysis because it is inconsistent. To overcome this drawback, Cucala (2006) defined the density of events locations as $\lambda_0(x) = \lambda(x)/m$, where $m = \int_W \lambda(x)dx$ is the expected number of events lying in W, and proved that its kernel estimator is consistent.

Kernel intensity estimation for spatial point processes has been usually conducted with scalar bandwidth parameters, which can be quite restrictive when dealing with anisotropic and highly inhomogeneous point processes. Following the propssal of Cucala (2006) to obtain consistent estimators, and the philosophy of bivariate kernel density estimation for bandwidth selection, Fuentes-Santos et al. (2016) propose a kernel estimator of the density of event locations with bandwidth matrix:

$$\hat{\lambda}_{0,H}(x) = \frac{\hat{\lambda}_H(x)}{N} 1\left(N > 0\right) = \left(p_H(x)N\right)^{-1} |H|^{-1/2} \sum_{i=1}^{N} k\left(H^{-1/2}\left(x - \mathbf{x_i}\right)\right) 1\left(N > 0\right) \tag{A-3}$$

where H is a symmetric and positive-definite matrix, and $|H|$ denotes the determinant of H. This banwidth matrix can be estimated using the plug-in algorithm detailed in Fuentes-Santos et al. (2016), which minimizes the asymptotic mean integrated square error

(AMISE) of $\hat{\lambda}_{0,H}(x)$

$$AMISE(H) = \frac{1}{4}\mu_2(k)^2 \int_{\mathbb{R}^2} tr\left(HD^2\lambda_0(x)\right)^2 dx + A(m)|H|^{-1/2}R(k) \tag{A-4}$$

where $\mu_2(k)^2 = |\int_{\mathbb{R}^2} uu^T k(u)du|$, $D^2\lambda_0(x)$ is the Hessian matrix of λ_0, $R(k) = \int_{\mathbb{R}^2} k(x)^2 dx$, $A(m) = E\left[\frac{1}{N}I[N>0]\right] = e^{-m}\sum_{k=1}^{\infty}\frac{m^k}{kk!} < e^{-m}\sum_{k=0}^{\infty}\frac{2m^k}{(k+1)!} = 2/m \to 0$, and for any matrix A, $tr(A)$ denotes the sum of its diagonal terms.

APPENDIX A.2. Nonparametric comparison of first-order intensity functions

A common question in the analysis of multitype spatial point processes is whether two types of events have the same spatial structure. In this application we can wonder, for instance, if gunfires with and without mortal victims have the same spatial distribution or there is any area with high mortality risk. Taking into account that the first-order intensity describes the distribution of events in the observation domain, comparing the first-order intensities of two point patterns appears as a natural way to address this question.

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 $\mathbf{X_1}$, $\mathbf{X_2}$, the spatial patterns of type 1 and type 2 events in $\mathbf{X}$. Although the Poisson assumption is required to guarantee the consistency of the kernel densities of event locations and to obtain the asymptotic null distribution of the test statistic introduced bellow. We can also apply this test to compare non-Poisson point processes. It should also be noted that the Poisson assumption is not very restrictive, as assuming that the point process is Poisson to estimate its first-order properties of an observed pattern is the common practice when additional information, such as covariates, is not available. If $\mathbf{X_1}$ and $\mathbf{X_2}$ have the same spatial distribution their first-order intensities, $\lambda_1(x)$ and $\lambda_2(x)$, are proportional and, consequently, they have the same density of event locations. 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, we can extend the nonparametric test developed by Duong et al. (2012) for multivariate data to the spatial point process framework and use an L^2-distance to measure the discrepancy between the density of event locations of $\mathbf{X_1}$ and $\mathbf{X_2}$

$$\begin{aligned}
T &= \int_W \left(\lambda_{01}(x) - \lambda_{02}(x)\right)^2 dx && \text{(A-5)} \\
&= \int_W \lambda_{01}(x)^2\, dx + \int_W \lambda_{02}(x)^2\, dx - \int_W \lambda_{01}(x)\,\lambda_{02}(x)\, dx - \int_W \lambda_{02}(x)\,\lambda_{01}(x)\, dx \\
&= \mathbb{E}_{\mathbf{X_1}}\left[\lambda_{01}(x)\right] + \mathbb{E}_{\mathbf{X_2}}\left[\lambda_{02}(x)\right] - \mathbb{E}_{\mathbf{X_2}}\left[\lambda_{01}(x)\right] - \mathbb{E}_{\mathbf{X_1}}\left[\lambda_{02}(x)\right]
\end{aligned}$$

Expression (A-5) can be rewritten as $T = \psi_1 + \psi_2 - (\psi_{12} + \psi_{21})$, where $\psi_j = \mathbb{E}_{\mathbf{X_j}}\left[\lambda_{0j}(x)\right]$ and $\psi_{ij} = \mathbb{E}_{\mathbf{X_i}}\left[\lambda_{0j}(x)\right]$, for $i, j = 1, 2$. Assuming $W = \mathbb{R}^2$ to avoid the limitation of edge-effects, and using kernel smoothing to estimate each component of T we obtain the test statistic

$$\hat{T} = \hat{\psi}_1 + \hat{\psi}_2 - \left(\hat{\psi}_{12} + \hat{\psi}_{21}\right) \tag{A-6}$$

where,

$$\begin{aligned}
\hat{\psi}_j &= N_j^{-2} \sum_{k_1=1}^{N_j} \sum_{k_2=1}^{N_j} k_{G_j}\left(\mathbf{x_{k_1}} - \mathbf{x_{k_2}}\right) I\left(N_j > 0\right) \;\; j = 1, 2 \\
\hat{\psi}_{ij} &= \left(N_i N_j\right)^{-1} \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} k_{G_i}\left(\mathbf{x_k} - \mathbf{x_l}\right) I\left(N_i > 0\right) I\left(N_j > 0\right) \;\; i, j = 1, 2
\end{aligned}$$

where G_1 and G_2 are the bandwidth matrices for the kernel estimators of the density functionals ψ_1 and ψ_2. Given the closeness between the kernel estimators of the density of event locations of spatial point processes and the density of bivariate random variables, we select these bandwidths using the plug-in algorithm proposed by Chacón and Duong (2010) for integrated density derivatives considering, in this case, order $r = 0$.

$$\hat{T} = \hat{\psi}_1 + \hat{\psi}_2 - \left(\hat{\psi}_{12} + \hat{\psi}_{21}\right) \tag{A-7}$$

The null distribution of $\hat{T}$ is asymptotically normal under regularity conditions analogous to those assumed in the classical multivariate distribution framework. However, the slow convergence rate to the normal distribution discourages using the asymptotic null distribution as calibration procedure. For this reason, Fuentes-Santos et al. (2017) developed a smooth bootstrap algorithm to calibrate the test, which is implemented as follows:

1. Compute the test statistic $\hat{T}_1$ for the observed patterns, $\mathbf{X_1}$ and $\mathbf{X_2}$.

2. Let $\mathbf{X} = \{\mathbf{X_1}, \mathbf{X_2}\}$ be the unmarked spatial point pattern comprising both types of events, obtain the kernel estimator of its density of event locations, $\hat{\lambda}_{0,H}(x)$.

3. For $b = 2, \ldots, B$:

   3.1 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 $\mathbf{X}_j$.

3.2 Compute the corresponding test statistic $\hat{T}_b$.

4. The probability of rejecting the null hypothesis is the proportion of bootstrap statistics $\{\hat{T}_b\}_b^B$ larger than $\hat{T}_1$.

Two bandwidth selection procedures are involved in the calibration algorithm. The bandwidth matrix for the kernel density of event locations (A-3) in step 2 is obtained by the plug-in bandwidth selector introduced by Fuentes-Santos et al. (2016). whereas, as stated above, we use the plug-in algorithm proposed by Chacón and Duong (2010) to select the bandwidth matrices, $G_j, j = 1, 2$, for the kernel estimators of the squared density integrals $\psi_j, j = 1, 2$ in (A-7) to conduct steps 1 and 3.2.

APPENDIX A.3. Kernel estimator and tolerance contour for the log-relative risk function

If the T-test finds differences between, for instance gunfire with and without mortal victims, we may be interested on where did those differences occur. To answer this question we can use the relative-risk function, introduced by Bithell (1990) to compare the geographical distribution of disease cases, $\mathbf{X}_1$, and a random sample of the population at risk referred as controls, $\mathbf{X}_2$. The relative risk function is defined as the ratio between the densities of event locations of cases and controls, $r(x) = \lambda_{01}(x)/\lambda_{02}(x)$. Given that $\lambda_{0j}(x), j = 1, 2$ are strictly positive, and taking into account that the number of cases is often much smaller than the number of controls, it is usual to work with the log-relative risk $\rho(x) = \log(\lambda_{01}(x)/\lambda_{02}(x))$ to handle this asymmetry.

Kelsall and Diggle (1995) proposed estimating the log-relative risk function as the ratio between the kernel density of event locations of cases and controls $\hat{\rho}(x) = \log\left(\hat{\lambda}_{01,h}(x)/\hat{\lambda}_{02,h}(x)\right)$, where $\hat{\lambda}_{0j,h}(x) = N_j^{-1}\hat{\lambda}_j(x)I\left(N > 0\right)$ with $\hat{\lambda}_j(x)$ as defined in expression (A-2). Kelsall and Diggle (1995) also developed a least-squares cross-validation bandwidth selector and proposed using the same bandwidth parameter $h = h_1 = h_2$ for cases and controls, which leads to bias cancellation in areas where $\lambda_{01} = \lambda_{02}$.

When, as in this case, we are interseted on local features, fixed bandwidth kernel estimators can lead to disappointing results, as a single bandwidth is not able to simultaneously provide sufficient degree of smoothing in areas with low density and capture details in those with high density. Davies and Hazelton (2010) pointed out this drawback and proposed using an adaptive kernel estimator to overcome this limitation. Therefore we replace $\hat{\lambda}_{0j}(x),\ j = 1, 2$ by

$$\hat{\lambda}_{0j,h_j}\left(x\right) = \frac{1}{N_j}\sum_{i=1}^{N_j}\frac{1}{p_{h_{ji}}(x)h_{ji}^2}k\left(\frac{x - \mathbf{x_{ji}}}{h_{ji}}\right)I\left(N_j > 0\right) \qquad \text{(A-8)}$$

where $h_{ji} = h(\mathbf{x_{ji}})$ is a local bandwidth, which adapts the degree of smoothing to the number of events in the neighborhood of $\mathbf{x_{ji}}$ and, consequenlty, gives more accurate estimates of local features than a fixed bandwidth parameter. Following the proposal of Abramson (1982), the adaptive bandwidth for $\lambda_{0j}(x)$ is defined as

$$h_{ji}(x) = \frac{h_{0j}}{\gamma_j\lambda_{0j}(x)} \qquad \text{(A-9)}$$

where h_{0j} is a fixed smoothing parameter referred as global bandwidth and $\gamma_j = \exp\{N_j^{-1}\sum_{i=1}^{N_j}\lambda_{0j}\left(\mathbf{x_{ji}}\right)\}$ is a normalizing parameter that reduces the dependence of the bandwidth on the scale of the data.

In practice, as the density of event locations is unknown, we cannot obtain the adaptive bandwidth directly. We need to replace $\lambda_{0j}(x)$ in (A-9) by a pilot estimator. This pliot estimator is usually obtained by kernel smoothing, which requires the selection of a pilot bandwidth, h_{pj}. Davies and Hazelton (2010) propose using optimal smoothing (Terrell, 1990) to select the global bandwidth h_{0j}, and $h_{pj} = 0.5h_{0j}$ as pilot bandwidth.

Following the proposal of Davies et al. (2016), we use a common bandwidth function $h(\mathbf{x}) = h_0 \left(\gamma\lambda_0(x)\right)^{-1}$, where $\lambda_0(x)$ is the density of event locations of the unmarked point processes, $\mathbf{X} = \mathbf{X_1} \cup \mathbf{X_2}$, to estimate the density of event locations of cases and controls. This symmetric bandwidth reduces the halo effect, which generates artificial high risk rings in the boundary of areas with high control density but low risk. In addition, given that the edge corrector $p_{h_{ji}}(x)$ depends on the spatial locations and the bandwidth parameter, this term cancels out in $\hat{\rho}(x)$ when we use the same bandwidth function for cases and controls.

Once estimated the log-relative risk function, we can generate Monte-Carlo tolerance countours to identify areas with high mortality risk or police presence. As proposed by Kelsall and Diggle (1995), the tolerance contours are computed through the following algorithm

1. Simulate $B-1$ pairs of inhomogeneous point patterns with expected number of events equal to those of the observed case and control patterns and first-order intensity proportional to that of the unmarked point processes $\mathbf{X} = \mathbf{X_1} \cup \mathbf{X_2}$.
2. Estimate the log-relative risk function for each case-control simulated pattern, $\{\hat{\rho}_b(x)\}_{b=2}^{B}$.

3. Compute the tolerance surface of significance level α, $\hat{\rho}_{1-\alpha}(x)$, as the $100\,(1-\alpha)$ percentile of the Monte-Carlo log-relative risks at each location.

4. We have high risk functions at any location, x, where $\hat{\rho}\,(x) > \hat{\rho}_{1-\alpha}(x)$.

APPENDIX A.4. Log-ratio based separability test

Let $\mathbf{S} = \{(\mathbf{x_i}, \mathbf{t_i}), i = 1, \ldots, N\} \subset W \times T \subset \mathbb{R}^2 \times \mathbb{R}^+$be a realization of a spatiotemporal point process observed on a bounded domain $W \times T \subset \mathbb{R}^2 \times \mathbb{R}^+$, the spatio-temporal intensity function is a natural extension of the first-order intensity function of a spatial point process (Diggle, 2013)

$$\lambda(x,t) = \lim_{|dx \times dt| \to 0} \left\{ \frac{\mathbb{E}\left[N(dx,dt)\right]}{|dx \times dt|} \right\} \tag{A-10}$$

where $N(dx, dt)$ represents the number of events in the volume $dx \times dt$, dx is an infinitesimal disc containing the location x, and dt is an infinitesimal interval around time t.

One of the first steps in the analysis of any observed pattern is testing whether the spatiotemporal intensity is separable, i.e., whether it can be expressed as the product of its spatial and temporal components $\lambda(x,t) = \lambda_1(x)\lambda_2(t)$. Under separability the ratio between the spatio-temporal and spatial intensities, $\rho(x,t) = log\,(\lambda(x,t)/\lambda_1(x))$, does not depend on the spatial locations, x, for any $t \in T$. Considering this property Fuentes-Santos et al. (2018) propose using a no effect test that checks whether the log-ratio function $\rho(x,t) = \lambda(x,t)/\lambda(x)$ depends on the spatial locations.

Let $\mathbf{S} = \{(\mathbf{x_i}, \mathbf{t_i}), i = 1, \ldots, N\} \subset W \times T \subset \mathbb{R}^2 \times \mathbb{R}^+$ be a realization of an inhomogeneous spatiotemporal Poisson point process, and $\mathbf{X} = \{\mathbf{x_i}, i = 1, \ldots, N\} \subset W \subset \mathbb{R}^2$ the corresponding marginal spatial point process, and let $\lambda_0(x,t) = \lambda(x,t)/m$ and $\lambda_{01}(x) = \lambda_1(x)/m$, where $m = \int_T \int_W \lambda(x,t)dxdt = \int_W \lambda_1(x)dx$ is the expected number of events of both the spatial and spatiotemporal point processes, be their densities of event locations. The ratio

$$r(x,t) = \frac{\lambda(x,t)}{\lambda_1(x)} = \frac{\lambda_0(x,t)}{\lambda_{01}(x)}$$

can be seen as a spatiotemporal relative risk function whose control distribution remains constant over time (Sarojinie Fernando and Hazelton, 2014). The log-ratio function $\rho(x,t) = log(\lambda_0(x,t)/\lambda_{01}(x))$ can be estimated by kernel smoothing

$$\hat{\rho}(x,t) = \log \frac{\hat{\lambda}_{0,h_s,h_t}(x,t) + \delta}{\hat{\lambda}_{01,h_s}(x) + \delta} = \log\left(\hat{\lambda}_{0,h_s,h_t}(x,t) + \delta\right) - \log\left(\hat{\lambda}_{01,h_s}(x) + \delta\right) \qquad \text{(A-11)}$$

where $\hat{\lambda}_{0,h_s,h_t}(x,t)$ and $\hat{\lambda}_{01,h_s}(x)$ are the kernel estimators of the spatiotemporal and spatial densities of event locations, respectively. $h_s = (h_{s1}, h_{s2})$ denotes the main diagonal of the common diagonal bandwidth matrix used in the spatial component of the kernel estimators of $\lambda_0(x,t)$ and in the kernel estimator of $\lambda_{01}(x)$, and h_t is the scalar bandwidth for the temporal component in the numerator of $\hat{\rho}(x,t)$. In this work the optimal bandwidth was selected by least-squares cross-validation (LSCV). δ is a stabilizing constant that reduces the negative effect of data sparseness on the log-ratio estimator.

For a separable spatiotemporal point process, $\lambda(x,t) = \lambda_1(x)\lambda_2(t)$, therefore the log-ratio function, $\rho(x,t) = log\left(\lambda(x,t)/\lambda_1(x)\right)$ does not depend on the spatial locations, x, for any $t \in T$. Thus we have a regression problem where the log-ratio function evaluated at each event, $Y = \{y_i = \rho(\mathbf{x_i}, \mathbf{t_i}), i = 1, \ldots, n\}$ is a response variable that may depend on the spatial covariate $X = \{x_i = (x_{i1}, x_{i2}), i = 1, \ldots, n\}$ comprising the event locations, and we test for the effect of X on Y. Following Bowman and Azzalini (1997), we shall discriminate between two competing models

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

where, for any $x = (x_1, x_2) \in \mathbb{R}^2$, $m(\cdot)$ is an unknown smooth function, which can be estimated by kernel regression (Nadaraya, 1964; Watson, 1964)

$$\hat{m(x)} = \hat{m}(x_1, x_2) = \frac{\sum_{i=1}^{n} w_{g_1}(x_{i1} - x_1) w_{g_2}(x_{i2} - x_2) y_i}{\sum_{i=1}^{n} w_{g_1}(x_{i1} - x_1) w_{g_2}(x_{i2} - x_2)} \tag{A-12}$$

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, CV, and (iii) an AICC-based method.

Once computed $\hat{y} = \sum_{i=1}^{n} y_i$, which is the empirical estimator of μ in $\mathcal{H}_0$, and the regression function, $\hat{m}(\cdot)$ in $\mathcal{H}_1$, we compute the residual sum of squares for the null, RSS_0, and alternative, RSS_1, models and define the generalized test statistic

$$F = \frac{(RSS_0 - RSS_1) \,/\, (df_1 - df_0)}{RSS_1/df_1} \tag{A-13}$$

where df_0 and df_1 denote the degrees of freedom for the residuals under each hypothesis. This separability test is refereed here as *F-test*.

Bowman and Azzalini (1997) proposed two calibration procedures: (i) a χ^2 approximation of the null distribution of F, which can be used when the errors of the regression model are normal, and (ii) a computationally intensive procedure based on permutation tests otherwise.

The permutation test relies on the fact that under $\mathcal{H}_0$ the pairing of any particular x and y is completely random. Then, the distribution of the test statistic, F, can be generated by simulation, using random pairings of the observed values of X and Y and computing the corresponding test statistic in each case. The empirical p-value of the test is the proportion of simulated F-statistics larger than that obtained from the observed data.

APPENDIX B. References

References

Abramson, I. S. (1982). On bandwidth variation in kernel estimates-a square root law. *The annals of Statistics*, 1217–1223.

Arias, E. D. and N. Barnes (2017). Crime and plural orders in Rio de Janeiro, Brazil. *Current Sociology 65*(3), 448–465.

Baddeley, A., J. Møller, and R. Waagepetersen (2000, November). Non- and semiparametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica 54*(3), 329–350.

Baddeley, A., E. Rubak, and R. Turner (2015). *Spatial point patterns: methodology and applications with R.* Chapman and Hall/CRC.

Besag, J. (1977). Contribution to the discussion of Dr Ripley's paper. *Journal of the Royal Statistical Society, Series B (Methodological) 39*(2), 193–195.

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

Bowers, K. J., S. D. Johnson, and K. Pease (2004). Prospective hot-spotting: The future of crime mapping? *British Journal of Criminology 44*(5), 641–658.

Bowman, A. W. and A. Azzalini (1997). *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*. Oxford Statistical Science Series 18.

Bowman, A. W. and A. Azzalini (2014). *R package* `sm`*: nonparametric smoothing methods (version 2.2-5.5)*. University of Glasgow, UK and Università di Padova, Italia.

Chacón, J. E. and T. Duong (2010). Multivariate plug-in bandwidth selection with unconstrained pilot bandwidth matrices. *Test 19*(2), 375–398.

Chainey, S., L. Tompson, and S. Uhlig (2008). The utility of hotspot mapping for predicting spatial patterns of crime. *Security Journal 21*(1-2), 4–28.

Cohen, J., W. L. Gorr, and A. M. Olligschlaeger (2007). Leading indicators and spatial interactions: A crime-forecasting model for proactive police deployment. *Geographical Analysis 39*(1), 105–127.

Cucala, L. (2006). *Espacements bidimensionnels et données entachés d'erreurs dans l'analyse des procesus ponctuels spatiaux*. Ph. D. thesis, Université des Sciences de Toulouse I.

Davies, T. M. and M. L. Hazelton (2010). Adaptive kernel estimation of spatial relative risk. *Statistics in Medicine 29*(23), 2423–2437.

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.

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), 1191–1221.

Diggle, P. (1985). A kernel method for smoothing point process data. *Journal of the Royal Statistical Society. Series C (Applied Statistics) 34*(2), pp. 138–147.

Diggle, P. (2013). *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns* (3rd ed.). Boca Raton, Florida: CRC Press.

Duong, T. (2018). *ks: Kernel Smoothing*. R package version 1.11.1.

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), 8382–8387.

Farrell, G. and K. Pease (2001). *Repeat victimization*, Volume 12. Criminal Justice Press.

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), 416–435.

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*, 240–260.

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.

Gabriel, E. and P. J. Diggle (2009). Second-order analysis of inhomogeneous spatio-temporal point process data. *Statistica Neerlandica 63*(1), 43–51.

Gerber, M. S. (2014). Predicting crime using Twitter and kernel density estimation. *Decision Support Systems 61*, 115–125.

Gorr, W. L. (2009). Forecast accuracy measures for exception reporting using receiver operating characteristic curves. *International Journal of Forecasting 25*(1), 48–61.

Johnson, S. D. (2008). Repeat burglary victimisation: a tale of two theories. *Journal of Experimental Criminology 4*(3), 215–240.

Johnson, S. D., W. Bernasco, K. J. Bowers, H. Elffers, J. Ratcliffe, G. Rengert, and M. Townsley (2007). Spacetime patterns of risk: a cross national assessment of residential burglary victimization. *Journal of Quantitative Criminology 23*(3), 201–219.

Kelsall, J. and P. Diggle (1995). Kernel estimation of relative risk. *Bernoulli 1*(1-2), 3–16.

Kennedy, L. W., J. M. Caplan, and E. Piza (2011). Risk clusters, hotspots, and spatial intelligence: risk terrain modeling as an algorithm for police resource allocation strategies. *Journal of Quantitative Criminology 27*(3), 339–362.

Liu, H. and D. E. Brown (2003). Criminal incident prediction using a point-pattern-based density model. *International Journal of Forecasting 19*(4), 603–622.

Loeffler, C. and S. Flaxman (2018). Is gun violence contagious? A spatiotemporal test. *Journal of Quantitative Criminology 34*(4), 999–1017.

Mohler, G. (2014). Marked point process hotspot maps for homicide and gun crime prediction in Chicago. *International Journal of Forecasting 30*(3), 491–497.

Mohler, G., M. B. Short, P. J. Brantingham, F. P. Schoenberg, and G. E. Tita (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association 106*(493), 100–108.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability & Its Applications 9*(1), 141–142.

Oliveira, R. E. S. (2018). Police and market: the creation of pacifying Police units in Rio de Janeiro. Master's thesis, University of Manitoba.

Passos, A. M. (2018). Fighting crime and maintaining order:shared worldviews of civilian and military elites in Brazil and Mexico. *Third World Quarterly 39*(2), 314–330.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Reinhart, A. and J. Greenhouse (2018). Self-exciting point processes with spatial covariates: modelling the dynamics of crime. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 67*(5), 1305–1329.

Ripley, B. (1977). Modelling spatial patterns (with discussion). *Journal of the Royal Statistical Society, Series B (Methodological) 39*, 172–212.

Ripley, B. (1981). *Spatial Statistics*. New York: Wiley.

Sarojinie Fernando, W. T. P. and M. L. Hazelton (2014, April). Generalizing the spatial relative risk function. *Spatial and Spatio-temporal Epidemiology 8*, 1–10.

Silva, L. J. S., S. F. Gonzáles, C. F. Almeida, S. D. Barbosa, and H. Lopes (2017). CrimeVis: An Interactive Visualization System for Analyzing Crime Data in the State of Rio de Janeiro.

Taddy, M. A. (2010). Autoregressive mixture models for dynamic spatial Poisson processes: Application to tracking intensity of violent crime. *Journal of the American Statistical Association 105*(492), 1403–1417.

Terrell, G. R. (1990). The maximal smoothing principle in density estimation. *Journal of the American Statistical Association 85*(410), 470–477.

Waagepetersen, R. P. (2007). An estimating function approach to inference for inhomogeneous neyman-scott processes. *Biometrics 63*(1), 252–258.

Wang, X., D. E. Brown, and M. S. Gerber (2012). Spatio-temporal modeling of criminal incidents using geographic, demographic, and Twitter-derived information. In *2012 IEEE International Conference on Intelligence and Security Informatics*, pp. 36–41.

Watson, G. S. (1964). Smooth regression analysis. *Sankhya: The Indian Journal of Statistics, Series A*, 359–372.

Weisburd, D., E. R. Groff, and S.-M. Yang (2012). *The criminology of place: Street segments and our understanding of the crime problem.* Oxford University Press.

Wooditch, A. and D. Weisburd (2016). Using spacetime analysis to evaluate criminal justice programs: An application to stop-question-frisk practices. *Journal of Quantitative Criminology 32*(2), 191–213.

Zhuang, J. and J. Mateu (2018). A semiparametric spatiotemporal Hawkes-type point process model with periodic background for crime data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, (avaiable online 04 December 2018).