Environ Ecol Stat (2009) 16:225–250
DOI 10.1007/s10651-007-0080-6

Modeling spatio-temporal wildfire ignition point patterns

Amanda S. Hering · Cynthia L. Bell · Marc G. Genton

Received: 1 March 2006 / Revised: 1 September 2006 / Published online: 8 March 2008

Abstract We analyze and model the structure of spatio-temporal wildfire ignitions in the St. Johns River Water Management District in northeastern Florida. Previous studies, based on the K-function and an assumption of homogeneity, have shown that wildfire events occur in clusters. We revisit this analysis based on an inhomogeneous K-function and argue that clustering is less important than initially thought. We also use K-cross functions to study multitype point patterns, both under homogeneity and inhomogeneity assumptions, and reach similar conclusions as above regarding the amount of clustering. Of particular interest is our finding that prescribed burns seem not to reduce significantly the occurrence of wildfires in the current or subsequent year over this large geographical region. Finally, we describe various point pattern models for the location of wildfires and investigate their adequacy by means of recent residual diagnostics.

Keywords Clustering · Diagnostics · Inhomogeneous · K-function · Marked · Multitype · Point process · Residuals

A. S. Hering · C. L. Bell · M. G. Genton (✉)
Department of Statistics, Texas A&M University, College Station, TX 77843-3143, USA
e-mail: genton@stat.tamu.edu

A. S. Hering
e-mail: mhering@stat.tamu.edu

C. L. Bell
e-mail: cynsimmons@stat.tamu.edu

M. G. Genton
Department of Econometrics, University of Geneva, Bd du Pont-d'Arve 40, 1211 Geneva 4, Switzerland
e-mail: Marc.genton@metri.unige.ch

1 Introduction

The spatio-temporal pattern of wildfire occurrences in the St. Johns River Water Management District (SJRWMD) of northeastern Florida is necessary information for state officials so that optimal firefighting resources and development projects can be placed in appropriate areas. A previous study by Genton et al. (2006) has analyzed this data using the second-moment K-function (Ripley 1977; Diggle 1983, p. 47) to determine whether this space-time point pattern shows a departure from complete spatial randomness. Complete spatial randomness (CSR), in this case, would indicate no significant spatial correlation between the wildfire occurrences, while a departure from CSR would indicate clustering or regularity of wildfire ignitions. One common assumption with this type of analysis (see also Podur et al. 2003) is that the point pattern of events has a constant intensity, λ, representing the mean number of events per unit area.

The study by Genton et al. (2006) used a dataset of reported wildfire occurrences in the SJRWMD that was collected by the Florida Division of Forestry from 1981 to 2001 containing the location (latitude and longitude of the centroid of the cadastral section) of the initial fire ignition, the date, the cause, the size of the fire in acres, and the fuel type (neither size nor fuel type were used in this study). In addition, the Florida Division of Forestry provided a GIS coverage which accurately maps 98% of reported locations of wildfire ignitions and acres burned. The SJRWMD covers 31,681 km^2 in northeastern Florida, see Fig. 1. All reported wildfire occurrences in this area were considered in the analysis except for those in the Ocala National Forest due to data constraints on federal lands. Over the 21 year period, there were 31,693 fire ignitions recorded.

The first objective of this paper is to examine the assumption of constant intensity and reanalyze the wildfire data from the SJRWMD region using an inhomogeneous version of the K-function. Specifically, we intend to determine whether the use of an intensity function that depends on locations alters the inference on departure from CSR reported by Genton et al. (2006). Upon reanalysis, we found that departures from CSR are less likely to occur when estimating the K-function from an inhomogeneous point process. This result leads us to question the homogeneity assumption with this data, as well as the common assumption of homogeneity, when testing for CSR of a point process. The behavior of the K-function under homogeneous and inhomogeneous estimates of the intensity is further evaluated with a small simulation study; when trend is present in a point pattern, the homogeneous K-function overstates the departure of the pattern from CSR. From the modeling point of view, this information is important because it suggests that a (parametric or nonparametric) trend should be included in a model for wildfire ignitions.

In addition to detection of clustering among points, the relationship between points of two types is investigated with the K-cross function. The causes of fires are divided into 4 categories: accidents, arson, lightning, and railroads, and are also marked by the year the fire occurred. A pair of marks is selected, such as arson and lightning, and the K-cross function is designed to detect whether or not there is clustering or inhibition between the types of points. In other words, do arson fires occur more or less commonly near lightning fires, or are arson fires distributed with no regard to

Fig. 1 St. Johns River Water Management District as a section of Florida (courtesy of the St. Johns River Water Management District)

the locations of lightning fires? From a temporal perspective, the K-cross function is applied to pairs of sequential years to determine whether or not high fire years are followed by low or high fire years. The K-cross function has both a version assuming constant intensity and an inhomogeneous counterpart. As with the K-function, the degree of clustering is greatly reduced when an inhomogeneous trend is assumed.

Under this assumption, no temporal patterns are detected, and only arson and accident fires show a tendency to cluster.

Prescribed burns are a common fire suppression management strategy and have been shown to be effective in localized areas. Data on the locations of prescribed burns between 1993 and 2001 was available, and the K-cross function can be used to evaluate the efficacy of prescribed burning on a large scale (both spatially and temporally). Prescribed burns for a given year are compared with wildfires in its year and the following year with the K-cross function; results show that for each pair of years, there is not significant evidence to reject CSR between the prescribed burn locations and the wildfire locations. This may have implications for use of prescribed burning except in specialized circumstances.

Finally, a few simple models for wildfire point patterns are built based on clues from the data. These models determine the form of the intensity estimators for the inhomogeneous K and K-cross functions. New residual diagnostics introduced by Baddeley et al. (2005) allow us to graphically compare the models to assess fit. These models may allow us to simulate and reproduce the patterns that are seen in the wildfire ignitions and will be building blocks for future models based on covariates such as humidity, temperature, rainfall, and fuel type, see Butry et al. (2008) for regression models based on such covariates.

The structure of the article is the following. In Sect. 2, we revisit the analysis of the SJRWMD wildfire data based on an inhomogeneous K-function and argue that clustering is less important than initially thought based on our simulation results. In Sect. 3, we use K-cross functions to study multitype point patterns, both under homogeneity and inhomogeneity assumptions, and reach similar conclusions as above regarding the amount of clustering. In particular, we find that prescribed burns seem not to reduce significantly the occurrence of wildfires in the current or subsequent year. In Sect. 4, we describe various point pattern models for the location of wildfires and investigate their adequacy by means of recent residual diagnostics. All numerical computations are carried out with the R statistical package `spatstat` (Baddeley and Turner 2005). We conclude in Sect. 5.

2 Inhomogeneous point patterns

The previous analysis of Genton et al. (2006) used the K-function as a measure of departure from CSR with three common assumptions: homogeneity (first-order intensity is constant), stationarity (second-order intensity depends on the distance between events, not the exact location), and isotropy (second-order intensity depends on the distance, not direction) of a point process.

2.1 Intensity

The first step in our analysis is to investigate the assumption of homogeneous intensity. We will use a chi-squared test of independence to determine if the intensity of points in one area differs significantly from the intensity of points in other areas. This analysis was done by dividing the SJRWMD region into subregions and then counting the

Table 1 Chi-squared test statistics based on 3 and 9 equally sized subregions of the SJRWMD

Fire cause	X^2 for 3 divisions	X^2 for 9 divisions
All	4,042.9	6,401.1
Accident	2,158.6	2,807.8
Arson	1,438.3	3,670.6
Lightning	1,569.3	2,096.5
Railroad	418.7	621.8

number of fire events within each subregion. Due to the irregular shape of the region, the divisions were made horizontally to maintain approximately equal sizing of each subregion. The chi-squared test statistic is

$$X^2 = \sum_{i=1}^{m} \frac{(N_i - \bar{N})^2}{\bar{N}}, \tag{1}$$

where m is the number of subregions, N_i is number of observed events in subregion number i and $\bar{N} = \frac{1}{m}\sum_{i=1}^{m} N_i$ is the mean or expected number of events in each subregion under the assumption of homogeneity (Diggle 1983, p. 33). We maintained the chi-squared test of independence condition that no more than 20% of the cells can have expected counts less than five by restricting the number of subregions to three and nine. The three subregion test was used to evaluate an overall homogeneity while the nine subregion test was used to evaluate more regional homogeneity. We calculated X^2 for each subregion type and for all of the data over the 21 years as well as for the data divided into fire causes, see Table 1. At a 99% level, all of these chi-squared test statistic values are extremely large compared to $\chi^2_{2,0.99} = 9.21$ and $\chi^2_{8,0.99} = 20.09$ indicating that we have significant evidence for an inhomogeneous intensity.

2.2 Inhomogeneous intensity

We will employ the same second-order statistic, the K-function, in our analysis but will only change the assumption of homogeneous intensity to include all possible intensities. Intuitively the K-function can be described as

$$K(r) = \lambda^{-1}\mathrm{E}\ (\text{number of events within distance } r \text{ of a randomly chosen event}), \tag{2}$$

where E represents the mathematical expectation and λ is the (constant) mean number of events per unit area.

By using the inhomogeneous K-function to reanalyze the data, we remove the assumption of an underlying homogeneous point process while still assuming isotropic stationarity. Intuitively, the inhomogeneous K-function has the same interpretation as the homogeneous K-function (2), except that the intensity of events is no longer constant but depends on the location of the events. It is defined as

$$K_{inhom}(r) = \frac{1}{\nu(B)} \mathrm{E}\left(\sum_{\mathbf{x}_i \in X \cap B} \sum_{\mathbf{x}_j \in X \setminus \{\mathbf{x}_i\}} \frac{I(\|\mathbf{x}_i - \mathbf{x}_j\| \le r)}{\lambda(\mathbf{x}_i)\lambda(\mathbf{x}_j)} \right), \tag{3}$$

where $B \in \mathcal{B}_0$, the class of bounded Borel sets in $\mathbb{R}^d$, $d \ge 1$, $\nu(B)$ represents the area of B, I is the indicator function, X is the set of all events $\mathbf{x}_i$ and $\mathbf{x}_j$, and r is the maximum considered distance between $\mathbf{x}_i$ and $\mathbf{x}_j$ (Baddeley et al. 2000). We can still interpret the K-function intuitively as in (2), but now the intensity is a function evaluated at both locations $\mathbf{x}_i$ and $\mathbf{x}_j$, that is, $\lambda(\mathbf{x}_i)$ represents the mean number of events occurring at location $\mathbf{x}_i$. We use the following estimator of the inhomogeneous K-function:

$$\hat{K}_{inhom}(r) = \frac{1}{\nu(D)} \sum_{\mathbf{x}_i \in X \cap D} \sum_{\mathbf{x}_j \in \{X \cap D\} \setminus \mathbf{x}_i} \frac{I(\|\mathbf{x}_i - \mathbf{x}_j\| \le r)}{\hat{\lambda}(\mathbf{x}_i)\hat{\lambda}(\mathbf{x}_j) w_{\mathbf{x}_i, \mathbf{x}_j}}, \tag{4}$$

where D is the complete domain of the dataset and $w_{\mathbf{x}_i, \mathbf{x}_j}$ is Ripley's edge correction factor.

Specifically, in both analyses, the L-function was employed since it is easier to visualize a departure from CSR. The L-function is defined as

$$L(r) = \sqrt{\frac{K(r)}{\pi}}. \tag{5}$$

By transforming the K-function in this manner, $L(r) = r$ under CSR. We can use this line through the origin as a reference to compare with our estimated values from the data. Additionally, we simulated 100 values of $L(r)$ under CSR with the estimated intensity function and used the minimum and maximum values to construct a confidence envelope around the line $L(r) = r$. If the estimated $L(r)$ falls outside this simulated confidence envelope then we reject the null hypothesis of CSR.

In order to compute $\hat{K}_{inhom}(r)$, we first need an accurate estimate of the intensity at each event location. Estimators of this nonconstant intensity to use in formula (4) can be obtained using a fitted parametric model or with nonparametric estimation. First, a model can be fitted to the data to describe both the "spatial trend" and "random interactions" of events. A class of parametric functions flexible enough to model the spatial trend for a given year or for a given cause in this dataset is the exponential of a fifth degree polynomial. Formally, the intensity can be modeled by

$$\lambda(\mathbf{x}) = \mathrm{e}^{\boldsymbol{\theta}^{\mathrm{T}} \mathbf{p}_{\mathbf{x}}}, \tag{6}$$

where $\mathbf{x} = (x, y)$, and $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_{21})^{\mathrm{T}}$ is a 21×1 vector of coefficients for the vector $\mathbf{p}_{\mathbf{x}} = (1, x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3, x^4, x^3y, x^2y^2, xy^3, y^4, x^5, x^4y, x^3y^2, x^2y^3, xy^4, y^5)^{\mathrm{T}}$. The coefficients $\boldsymbol{\theta}$ are obtained from the maximum pseudolikelihood model-fitting procedure in `spatstat`, and the corresponding intensity estimator for the points of each type can be used in (4).

Nonparametric estimation of the intensities (Diggle 1985; Møller and Waagepetersen 2003, p. 36), is another option for substitution into (4). A nonparametric kernel estimator of the intensity $\lambda(\mathbf{x})$ with bandwidth b over the domain A is

$$\hat{\lambda}_b(\mathbf{x}) = \frac{1}{p_b(\mathbf{x})} \sum_{i=1}^{n} \frac{1}{b^2} k\left(\frac{\mathbf{x} - \mathbf{x}_i}{b}\right), \tag{7}$$

where $p_b(\mathbf{x}) = \int_A b^{-2} k((\mathbf{x} - \mathbf{u})/b) d\mathbf{u}$ serves as the edge correction. The choice of the kernel, k, is not as important as the chosen bandwidth; we use the Gaussian kernel with scale parameter σ acting as the bandwidth. Too large a choice of σ reduces the estimate to the one for constant intensity; whereas, too small a choice of σ will capture local trends instead of the global trend.

With so many choices for intensity estimates, it is not obvious how to choose a "good" one. We fitted many parametric models and nonparametric models. Evaluating the trend and interaction components of a point process will be discussed in Sect. 4, and justification of the models chosen herein will be given.

2.3 Comparison of the results

The dataset consists of 31,693 wildfire ignition observations from 1981 to 2001 which is divided into the four causes of fires: accidents, arson, lightning, and railroads. After this division there are 14,535 fires due to accidents, 9,351 fires due to arson, 7,134 fires due to lightning, and 673 fires due to railroads. Because railroad fires are too few for some of the estimation procedures, we will only consider fires caused by accident, arson, and lightning. Figure 2 gives a representation of the density of all of the fire locations for the region and also plots each of accident, arson, and lightning caused fires alone. The sizes of the datasets by cause are still quite large, and for estimation purposes we must divide accident, arson, and lightning fires further into year categories from 1981 to 2001 creating 63 separate datasets. Even after this partitioning the datasets consist of a large number of wildfire incidents that must all be simultaneously used in the computations of L-functions and intensity estimates. Using an approximation to the homogeneous L-function based on the Fast Fourier Transform does reduce the size of the computations enough to handle datasets as large as accident, arson, and lightning (see Genton et al. 2006); however, no such equivalent approximation exists for the inhomogeneous case. Due to computational difficulties in calculating the L-function with nonparametric intensity estimators, the results for the estimated L-function are all computed using an intensity function in the form of an exponential of a fifth degree polynomial depending on the spatial coordinates.

In addition to computing the inhomogeneous L-functions for the data, we replicated the previous analysis based on the homogeneity assumption. These replications were used as standards to which to compare results. Due to the large number of datasets, we choose only to show the comparisons with arson fires in 1996 and 1997, with lightning fires in 1988 and 1997, and with accident fires in 1988 and 1997, see Figs. 3–5. All the datasets follow a similar pattern and are well represented by the behavior seen in these examples.

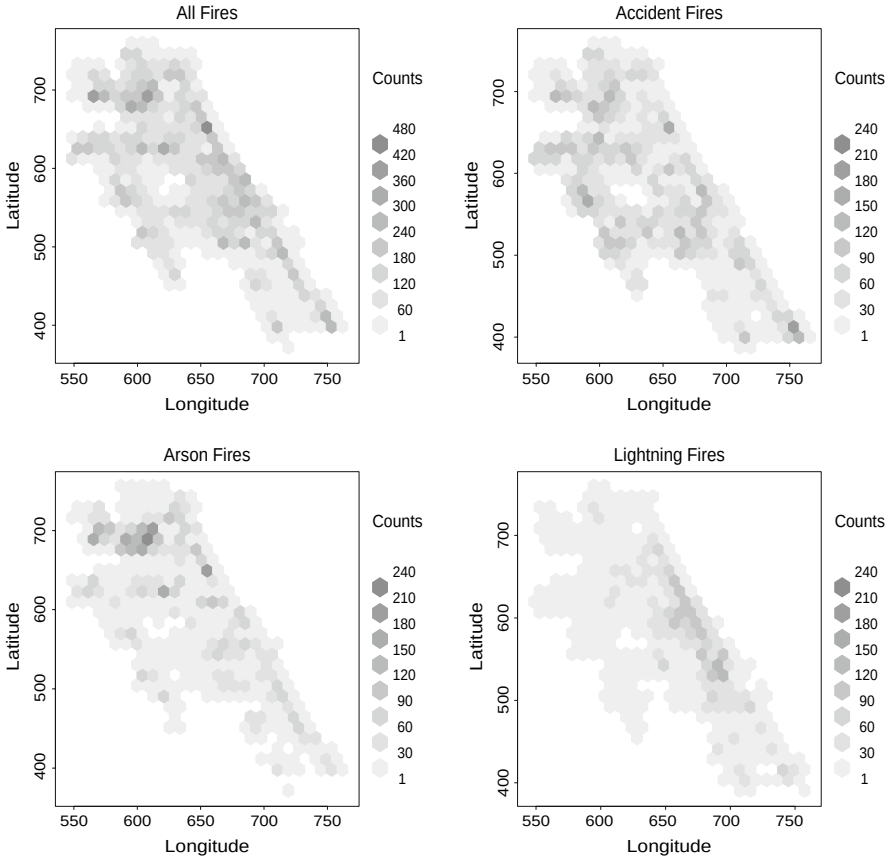

Fig. 2 Locations of all reported fires, and all reported fires sorted by accident, arson, or lightning caused fires in the St. Johns River Water Management District from 1981 to 2001

Two phenomena are apparent in comparing the homogenous L-function with the inhomogeneous L-function. The confidence envelopes widen for the inhomogeneous L-function, an indicator of greater variability in the intensity estimate. This also increases the region in which the data would support the null hypothesis of complete spatial randomness. In Fig. 3, this widening is especially apparent for small values of the radius r. Note also in Fig. 4 that only 135 lightning fires occurred in 1997, as opposed to 313 in 1988, accounting for the more erratic confidence envelopes in 1997. Secondly, the values of the L-function over almost all values of r decrease for the inhomogeneous estimate. This indicates that the homogeneous L-function appears to overestimate the interaction between the points when the trend across the dataset has not been taken into account.

For arson caused wildfires in 1996, the homogeneous L-function indicates clustering in these points up through 25 km, but the inhomogeneous L-function flirts with the upper bound through 15 km and then remains near the reference line thereafter. A similar pattern is exhibited in 1997, except that the strength of the clustering appears much stronger in the homogeneous case, and then the observed L-function in the

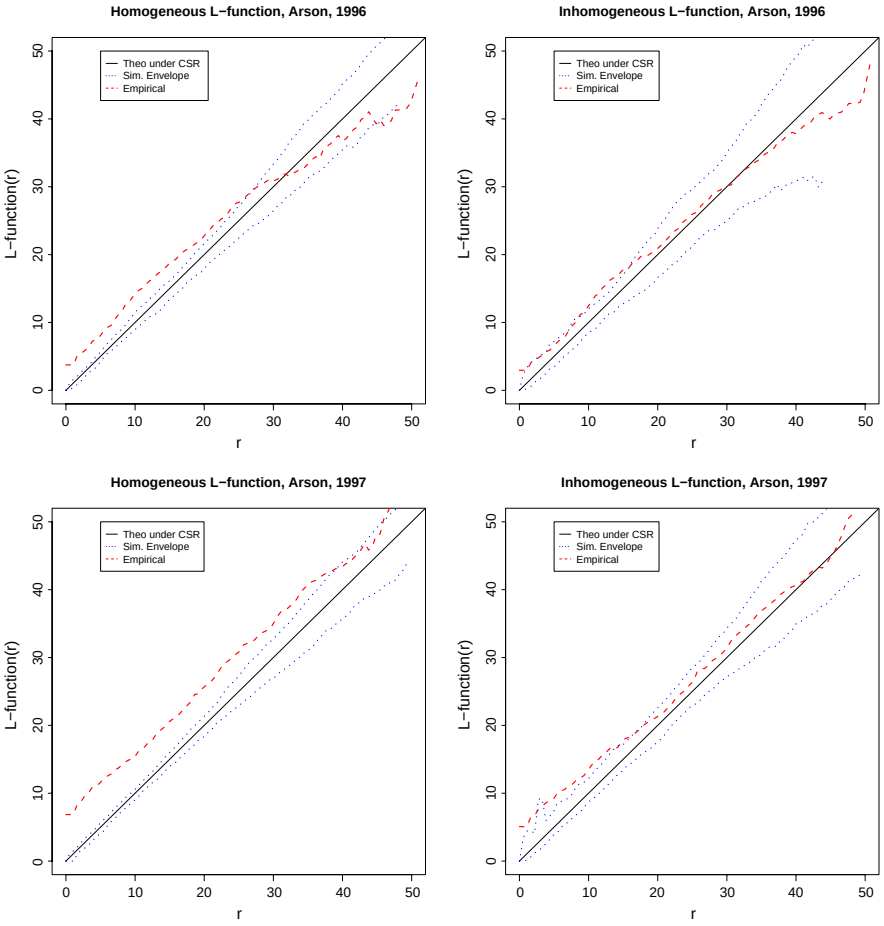

Fig. 3 Homogeneous and inhomogeneous L-function for arson wildfires, 1996 and 1997

inhomogeneous case is clearly above the upper bound through 10 km. In both years, significant clustering is detected within the arson fires, even given the adjustment for the spatial trend.

When the intensity of lightning fires is modeled inhomogeneously, the common finding that lightning fires are clustered (as in Genton et al. 2006) is called into question. Figure 4 shows that in both 1988 and 1997, while the homogeneous L-function detects significant clustering, the inhomogeneous L-function does not lie outside of the simulation envelopes. Therefore, even though a location may have an abundance of lightning fires, they do not appear to be significantly clustered together. Most lightning fires occur within a few miles of the coastline (see Fig. 2). This new finding suggests that incorporating a spatial trend is sufficient for describing the relationship among lightning wildfires. What has previously been interpreted as clustering among lightning fires may truly be due to the general ecological conditions, not interaction among the fires.

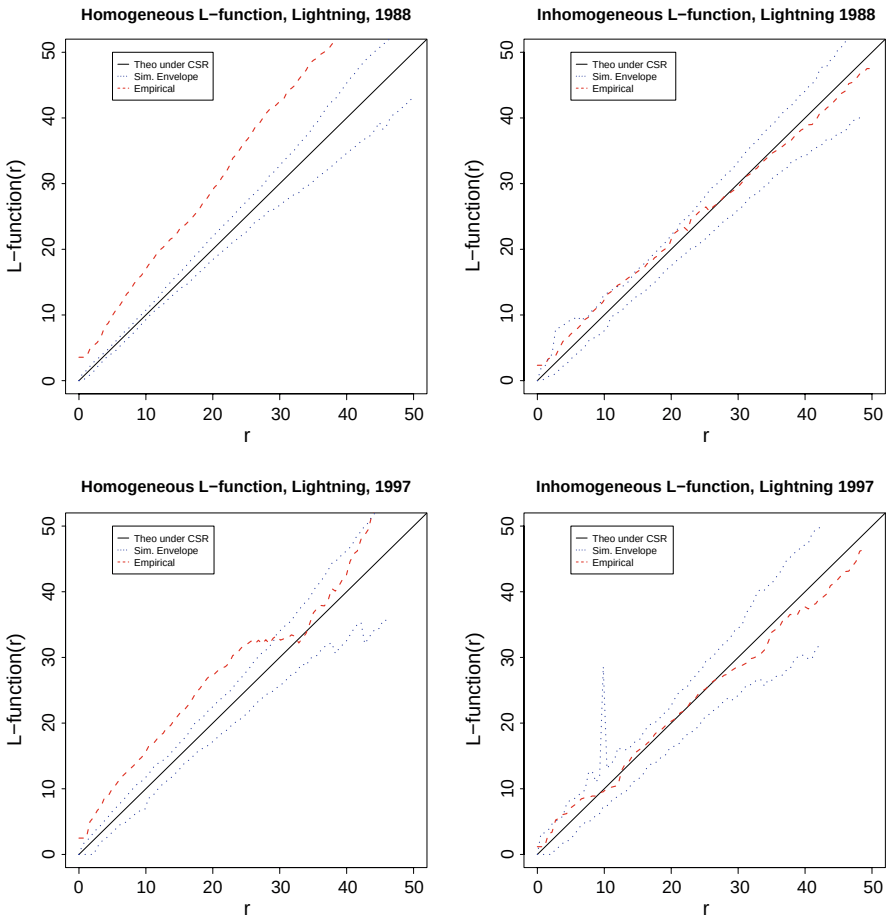

Fig. 4 Homogeneous and inhomogeneous L-function for lightning wildfires, 1988 and 1997

Accidental wildfires (Fig. 5) are the largest category containing 14,535 observations. Due to the nature of accidents being haphazard, this category accounts for many different types of fires, such as camping fires, equipment fires, and fires caused by children. All of these types of fires should be clustered around areas where people work, live, and recreate, and both homogeneous and inhomogeneous estimates of the L-function indicate some significant clustering, although it is reduced under the inhomogeneous intensity estimates. Like arson fires, this is another type of fire where significant clustering is observed even after adjusting for the trend in the data.

2.4 Simulation study

Throughout the previous analysis, we have stated that the homogeneous L-function is overestimating the amount of clustering present in the point pattern. Clustering and

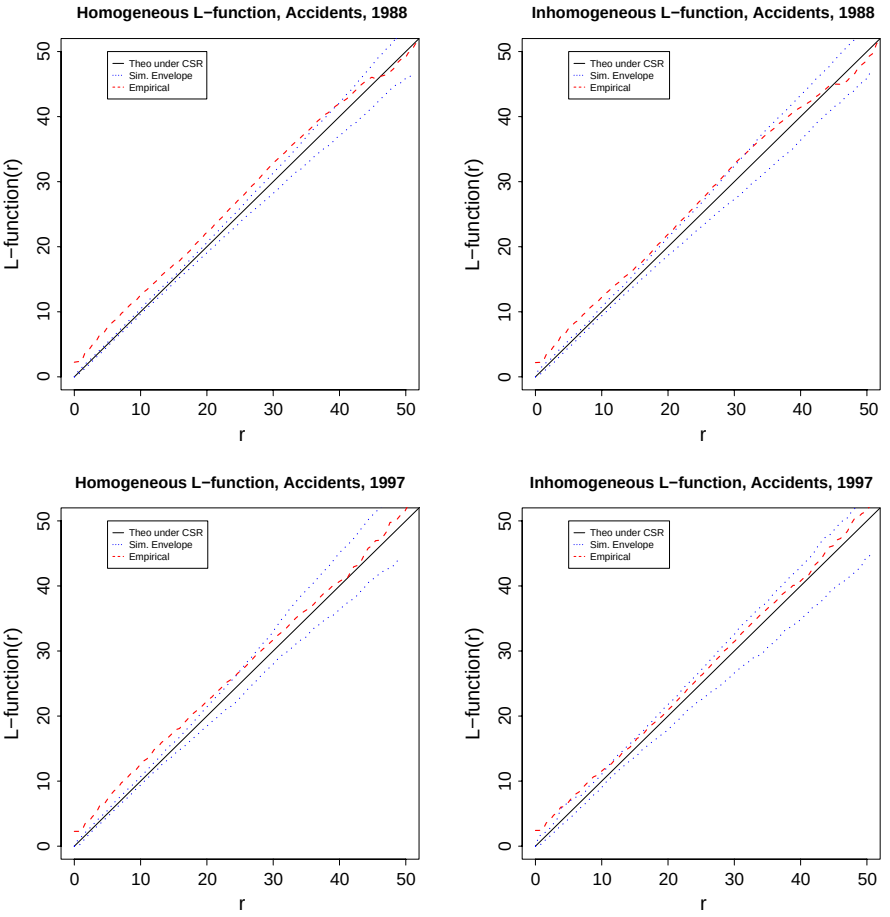

Fig. 5 Homogeneous and inhomogeneous L-function for accident wildfires, 1988 and 1997

interaction are two types of interpoint interaction, both of which the L-function is designed to detect. However, when a trend is present in the data as well as interaction, the inhomogeneous L-function seems to be better at detecting the interaction than the homogeneous L-function is. The wildfire locations in this dataset do display trend, as detected by the chi-squared test in Sect. 2.1. Therefore, we present the results of a simulation study to show the need to model the trend component of a point pattern in order to find interaction between the points.

Three types of point patterns are simulated: Geyer saturation process (no trend with interaction), inhomogeneous Strauss function with a strong, negative association (trend with interaction), and inhomogeneous Poisson process (trend without interaction). The specific trend functions and parameter specifications are taken from Baddeley et al. (2005). A realization from each of these point processes was generated on a unit square. In each case, both the homogeneous and inhomogeneous L-function was computed to test for departure from complete spatial randomness. For the homogeneous L-function,

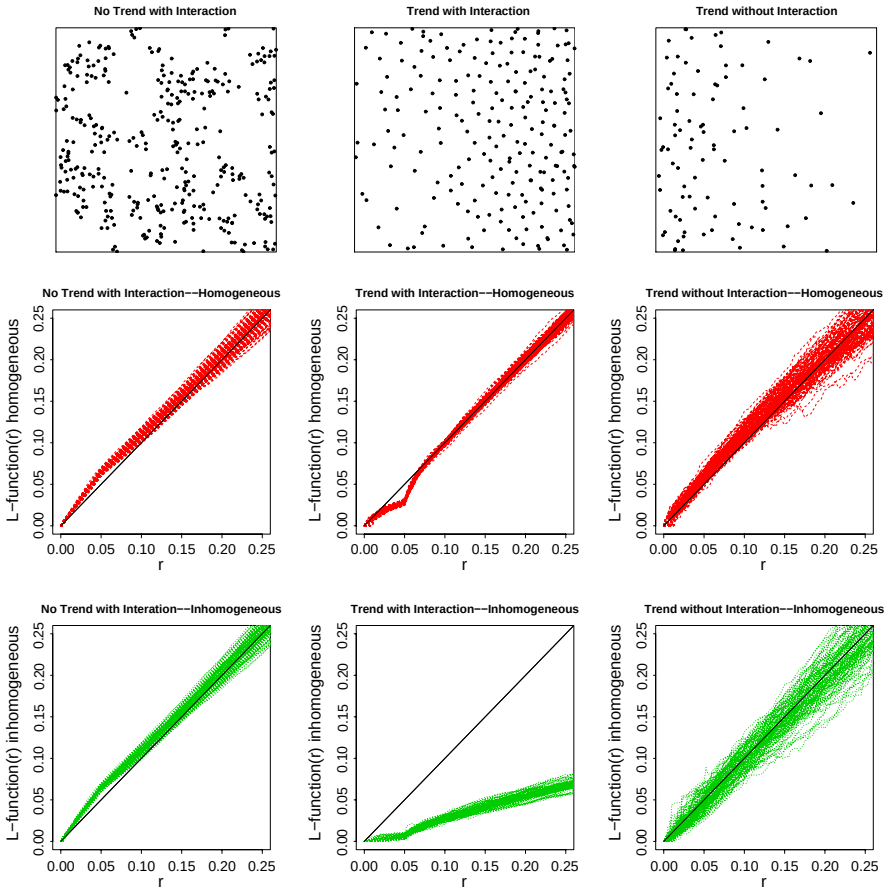

Fig. 6 The top panels represent a single realization of a point pattern with given trend and interaction components. The middle panels are the homogeneous intensity L-functions computed for each of 100 simulated datasets. The bottom panels are the inhomogeneous intensity L-functions computed for each of the same 100 simulated datasets

we assumed a constant intensity across the entire region. In the inhomogeneous case, we computed the intensity at each location using the trend function with which the point process was simulated. This process of simulating a realization of the point process and testing deviation from CSR was completed 100 times. A single realization of each of these three processes is given in the top panels of Fig. 6. The corresponding homogeneous and inhomogeneous L-functions for all 100 simulations are plotted underneath its point pattern.

When no trend is present, the intensity estimate for the inhomogeneous L-function is the same as the constant intensity used for the homogeneous L-function, so nothing is gained by using the inhomogeneous L-function in this case. Both functions detect the clustering at small distances r. However, when a trend is present, there are differences between the homogeneous and inhomogeneous L-function. With interaction as well as trend (middle column of panels in Fig. 6), the homogeneous L-function detects some inhibition at small distances, but the inhomogeneous L-function exhibits a much

greater degree of inhibition over all distances. But what is of greatest interest to us are the results in the final column of panels in Fig. 6. The L-function is computed here on datasets with trend but no interaction. In testing for clustering or inhibition, the null hypothesis is that there is no interaction, so this is what we should assume. When there truly is no interaction in these particular datasets, the homogeneous L-function is consistently greater than the inhomogeneous L-function. This indicates that the homogeneous L-function is misinterpreting the trend as clustering. Therefore, when trend is present in a dataset, incorporating this trend into the intensity estimation is important to accurately detect interpoint interaction.

3 Multitype point patterns

The locations of the wildfires comprise a spatial point pattern that is marked by both year and cause. Spatial interaction between points of two types occurs when events of each type are either closer or farther away than expected under the assumption that the two processes are independent. Questions such as "Are fires occurring in 1995 clustered around fires that occurred in 1994, or are they farther apart?" and "Do more accident caused fires occur nearby arson fires?" can be answered by analyzing the data with a function called K-cross and its relatives. Two versions of K-cross are applied to the SJRWMD data and compared—one assuming the process of each type is homogeneous and another assuming two inhomogeneous point processes.

3.1 K-cross function

The K-cross function is an extention of the K-function and includes information on the marks associated with each location. To begin, the point pattern is assumed to have a constant intensity for all of the points of both types. Proposed first by Ripley (1981), the K-cross function is

$$K_{ij}(r) = \lambda_j^{-1}\mathrm{E}\,(\text{number of type } j \text{ events within distance } r \text{ of a randomly chosen type } i \text{ event}), \tag{8}$$

where λ_j is the (constant) intensity of the marginal pattern of type j. If the bivariate spatial process is also stationary, in addition to homogeneous, then $K_{ij} = K_{ji}$.

The estimator of K_{ij} is a moment based estimator. Using notation borrowed from Schabenberger and Gotway (2005, p. 104), let the events of type i in a circle of area A, denoted $\nu(A)$, be observed with intensity λ_i and be referenced by the location $\mathbf{x}$. Events of type j in A then have intensity λ_j with location $\mathbf{y}$. Ripley's bias corrected estimator of K_{ij} is then

$$\hat{K}_{ij}(r) = \frac{1}{\hat{\lambda}_j} \cdot \frac{1}{\hat{\lambda}_i \nu(A)} \sum_k \sum_l \frac{I(\|\mathbf{x}_k - \mathbf{y}_l\| \le r)}{w_{\mathbf{x}_k,\mathbf{y}_l}}, \tag{9}$$

where $w_{\mathbf{x}_k,\mathbf{y}_l}$ is an edge correction representing the proportion of the circumference of the circle centered at location $\mathbf{x}_k$ with radius $\|\mathbf{x}_k - \mathbf{y}_l\|$ that lies inside of the region A. The estimator of the intensity for points of each type is $\hat{\lambda}_j = \frac{n_j}{\nu(D)}$, where n_j is the number of events of type j in the entire domain D. This is the maximum likelihood unbiased estimate when the process is a homogeneous Poisson process. The package `spatstat` has a function available to compute this estimator of $K_{ij}(r)$.

When the events of type i are independent of the events of type j, $K_{ij}(r)$ reduces to πr^2, regardless of the pattern of either type of event. Instead of plotting $\hat{K}_{ij}(r)$, it is then common to estimate and plot the L-cross function, defined as

$$L_{ij}(r) = \sqrt{\frac{K_{ij}(r)}{\pi}}. \tag{10}$$

The L-cross function has the nice property that under independence of the types of points, $L_{ij}(r) = r$, which is the reference line through the origin. Thus, it is straightforward to compare the observed data to what is expected under the null hypothesis. For example, if $L_{ij}(20) = 15$, then there are 5 fewer points of type j at a radius of 20 units from a randomly selected point of type i than would be expected if the two types of processes were completely independent of each other. Values of $L_{ij}(r)$ less than r indicate inhibition between the two types of points, and values greater than r indicate clustering between the two types of points.

For ease of comparison, both an L-index and simulation envelopes are computed. The L-index, defined by Genton et al. (2006) for the case of a homogeneous L-function, is an approximation of the area between $\hat{L}_{ij}(r)$ and the reference line. This is done by summing $\hat{L}_{ij}(r) - r$ over a range of values of r up to 45 km. This L-index will be particularly useful in comparing the change in the area between $\hat{L}_{ij}(r)$ and r over pairs of sequential years. In addition, simulation envelopes are formed using estimates of the intensity of events of type i, $\hat{\lambda}_i$, and of type j, $\hat{\lambda}_j$, assuming that each process is a homogeneous Poisson process. A set of points using each of the estimated intensities is simulated. Points simulated using $\hat{\lambda}_i$ are marked as type i points, and points simulated using $\hat{\lambda}_j$ are marked as type j points. These points are combined to create a single bivariate marked dataset. One hundred such datasets are simulated, $\hat{L}_{ij}(r)$ is computed for each dataset, and the maximum and minimum $\hat{L}_{ij}(r)$ out of all 100 datasets are taken to be the upper and lower bounds of the envelope. If $\hat{L}_{ij}(r)$ from the observed data crosses outside the upper or lower bounds, then the relationship between the types of points is considered to be significantly clustered or inhibited, respectively.

The one assumption that has already been seen to be violated is the assumption that the points are homogeneously distributed throughout the domain. In fact, similarly to the K-function, the significance of the interaction between types of points can be inflated when this assumption is not taken into consideration. The intensity estimators now become a function of location within the domain, $\hat{\lambda}_i(\mathbf{x})$ and $\hat{\lambda}_j(\mathbf{y})$, where $\mathbf{x}$ and $\mathbf{y}$ are any bidimensional locations in the domain. The `spatstat` package will compute $\hat{L}_{ij}(r)$ allowing for different estimators of intensity. The simulated datasets used to create the simulation envelopes must also be updated to adapt to the removal of this assumption. Now, the nonconstant intensity for the points of type i and type j

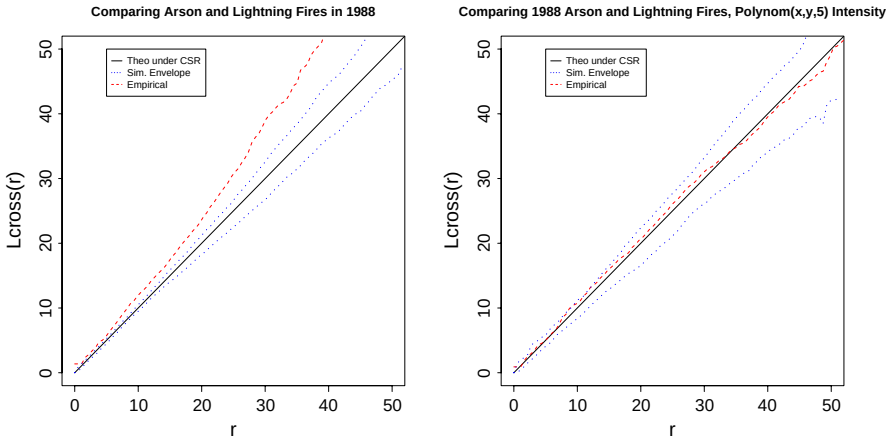

Fig. 7 L-cross function plotted for arson versus lightning fires in 1988 with homogeneous (left) and inhomogeneous, parametric (right) intensity estimators

are estimated, and points are simulated with these estimates and marked accordingly. Points simulated using $\hat{\lambda}_i(\mathbf{x})$ and points simulated using $\hat{\lambda}_j(\mathbf{y})$ are combined into a single dataset, and this procedure is repeated to create 100 datasets. Now, $\hat{L}_{ij}(r)$ is computed for each dataset using inhomogeneous estimates of the intensity of each dataset, and the maximum and minimum $\hat{L}_{ij}(r)$ over the 100 simulated datasets are the upper and lower bounds of the simulation envelope.

3.2 Comparison of the results

Because of the computational challenge of some of the calculations, interactions between causes were evaluated only for the years 1988 and 1997. The number of railroad caused fires during those two years was 58 and 6, respectively, which are too few to analyze. Instead, just comparisons between arson, accident, and lightning caused fires are made for the causes. In addition, temporal patterns are explored by examining the interaction between fires occurring in different years.

In examining the causes, each pair showed significant clustering using the constant intensity estimator, but most of the pairs had insignificant interaction with the parametric and nonparametric intensity estimators. Figure 7 demonstrates this situation. In the left panel using the constant intensity, as the radius around a randomly chosen arson caused fire increases, the number of fires caused by lightning within that circle is also increasing. The observed L-cross function is outside of the simulation envelope for almost every value of r, making it seem as if there is clustering between fires caused by lighting and those caused by arson. However, once the intensity estimation is changed to the parametric form in (6), the L-cross function (in the right panel) stays entirely inside of the simulation envelope. The interpretation here is that after adjusting for spatial trend, arson caused fires and lightning caused fires are distributed across the domain completely independently of each other. Given a particular location, there is no significant interaction between arson and lightning caused fires. The L-cross

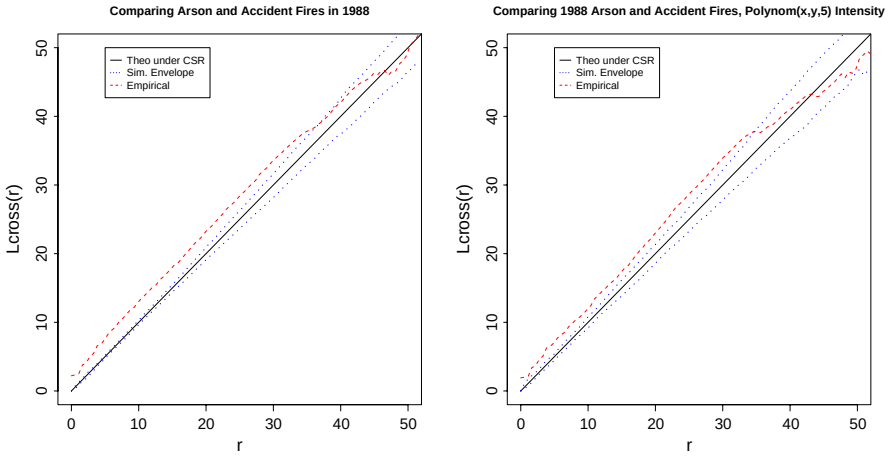

Fig. 8 L-cross function plotted for arson versus accident fires in 1988 with homogeneous (left) and inhomogeneous, parametric (right) intensity estimators

function plotted with the nonparametric intensity estimator yields the same result. Since lightning is a natural phenomenon, and arson is a man-made phenomenon, it may be expected that they should act independently of each other.

However, there is one comparison between the three causes that remains significant regardless of the intensity estimator. Between arson fires and accident fires there is significant clustering in 1988 as seen in Fig. 8 up to around 35 km. Thus, for shorter distances, more accident fires occur near arson fires and vice versa. Since accidents and arson are both caused by humans, we may expect them to both be occurring in or near locations where people recreate and live. Using nonconstant estimates of the intensities does not change this conclusion.

Genton et al. (2006) observed that a sequence of years with many fire events seemed to be followed by years with a low number of fire events. In this exploratory stage of the data analysis, it would then be reasonable to examine pairs of years over time to look for patterns. Pairs of years compared were those 1 year apart (81–82, 82–83, etc.); 2 years apart (81–83, 82–84, etc.); 3 years apart (81–84, 82–85, etc.); 4 years apart (81–85, 82–86, etc.); 5 years apart (81–86, 82–87, etc.); and 6 years apart (81–87, 82–88, etc.). Naturally, the sheer number of comparisons prohibits displaying each plot of $\hat{L}_{ij}(r)$ with its simulation envelope. However, the plots of L-index in Fig. 9 for each pair of years is still informative and much more concise. The solid line connects the L-index for homogeneous intensity, and the dashed line connects the values using the parametric intensity. The vertical lines connect two values, the L-indices computed for the upper and lower bound of the simulation envelope. If the observed L-index is inside of this bound, no or little interaction was evident in the plot of $\hat{L}_{ij}(r)$, but if the L-index is above (below) the bound, significant clustering (inhibition) is present.

It is obvious first that $\hat{L}_{ij}(r)$ computed with constant intensity appears to consistently overstate the significance of the clustering between the pairs of years. With the parametric intensity estimator, the L-index roughly mirrors that of the constant intensity but is much smaller and stays within its bounds. There does appear to be a consistent positive trend, but no significant interaction between fires occurring in

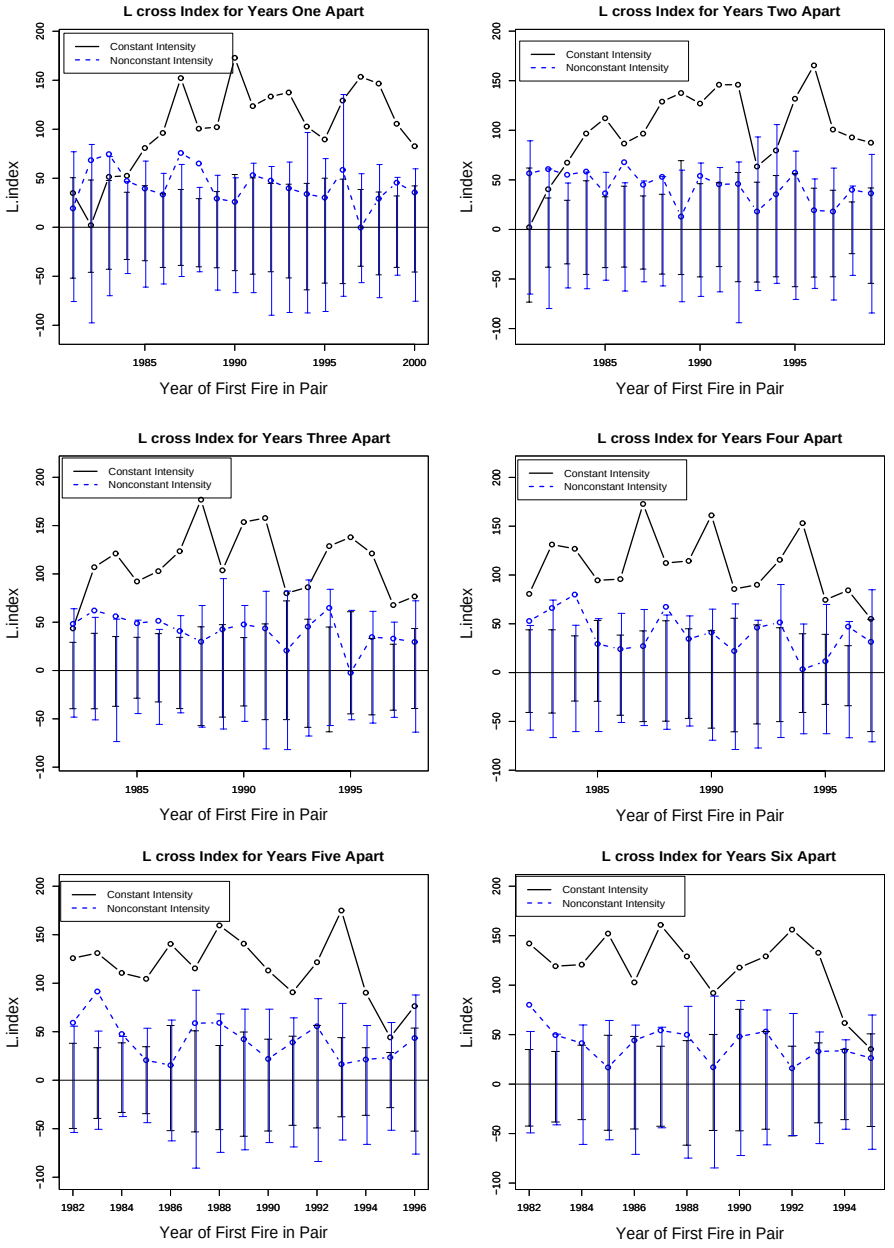

Fig. 9 *L*-index plotted for both homogeneous Poisson intensity and parametric intensity estimates. The vertical lines represent the corresponding *L*-index for the upper and lower simulation envelopes, and each year listed on the *x*-axis represents the first year in the pair of years being compared

1 year and those in subsequent years is evident. Thus, this analysis suggests that locations of fires in a given year would not be a good indicator of locations of fires in subsequent years.

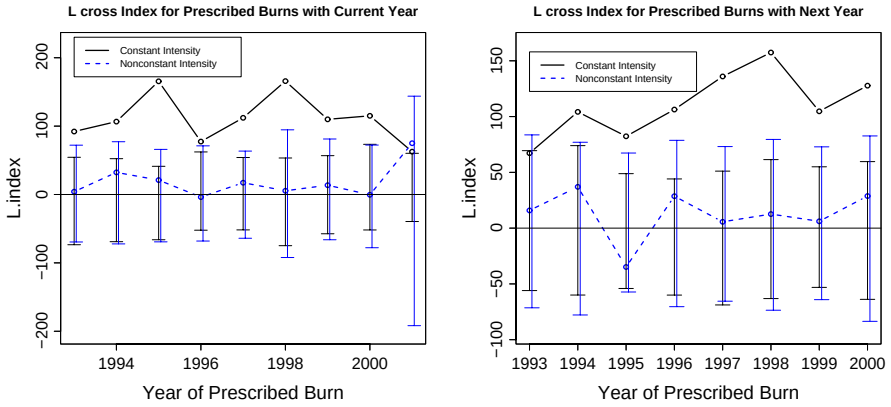

Fig. 10 *L*-index plotted for both homogeneous Poisson intensity and parametric intensity estimates. The left plot compares prescribed burns in the year on the horizontal axis with all wildfires the same year. The right plot compares prescribed burns with wildfires occurring the following year

3.3 Prescribed burns

Additional information on the locations of prescribed burns between the years 1993 and 2001 was available. The effects of prescribed burning have been widely studied in Florida. Prestemon et al. (2002) show that the kind of prescribed burn employed affects the wildfire risk differently with traditional burns either positively related to the current year's wildfire risk or unrelated. Brose and Wade (2002) report that prescribed burns are a short-term fix for suppressing wildfire since the vegetation reestablishes itself so quickly, and sometimes prescribed burns are actually fodder for surface fuels since they kill understory trees (Agee 2003).

Certainly, reduction in the number of wildfires is not the only goal land managers hope to achieve with prescribed burns (Haines et al. 2001; Outcalt and Wade 2004), but here, fire locations are given at a fine spatial scale (the section level), and the overall spatial interaction between all prescribed burns and all wildfires can be evaluated with the K-cross function. If the prescribed burns were significantly reducing the number of wildfires on a large scale, then the L-cross function and, thereby, L-index, would be expected to be negative and below the lower bound. Figure 10 shows that under the parametric intensity estimator, no significant interaction is occurring between the locations of prescribed burns and wildfires of the current or subsequent year.

Just because no significant interaction is occurring here does not mean that prescribed burns are useless or ineffective. On a small scale, they can be effective in preventing wildfires, but this analysis suggests that for a given location, prescribed burning does not reduce significantly the number of wildfires occurring in this region of Florida. Figure 11 shows the regions where prescribed burns and wildfires from 1993 to 2001 were most commonly located. Naturally, prescribed burns are implemented in areas where the risk of wildfire is the greatest, and we see that the greatest densities of prescribed burns are also where the highest densities of wildfires occur. If the prescribed burns that we have locations for are being used for fire suppression, then in a given location, we would expect fewer wildfire occurrences nearby where

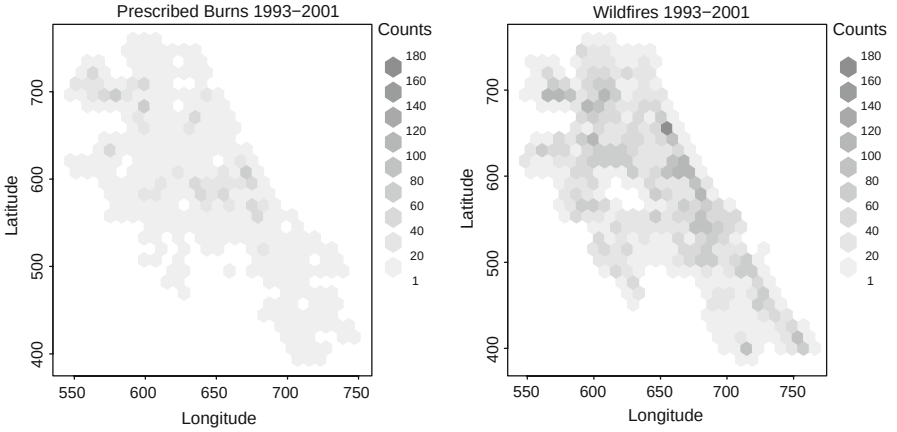

Fig. 11 Locations of all reported prescribed burns and all reported wildfires for the years 1993 through 2001

prescribed burns are conducted. In future analyses, the K-cross function could be implemented in subregions where prescribed burning is a common practice to further investigate its efficacy.

4 Point patterns modeling

The importance of using inhomogeneous estimates of intensity in the K and K-cross functions in order to accurately represent the amount of clustering was suggested by the simulation study in Sect. 2.4. Both parametric and nonparametric intensity estimates were discussed in Sect. 2.2. These estimates are based on the trend and interaction components of a model for a spatial point process. The trend is the spatial trend or systematic component of the model; the interaction describes the interpoint interaction structure (not to be confused with interaction between different types of points) or distributional component. For example, a Poisson process (homogeneous or inhomogeneous) would represent a point process with no interpoint interaction and independence between its points.

The goal of this section is twofold. First, we will discuss the recent residual diagnostics proposed by Baddeley et al. (2005) and summarize how these are used to determine the adequacy of a proposed model's fit. Then, the procedure for selecting the exponential of a fifth degree polynomial for the parametric intensity estimate and the bandwidth for the nonparametric intensity estimate will be described and justified with the residual diagnostics.

4.1 Model theory

The Pearson residuals defined by Baddeley et al. (2005) are

$$R\left(A, \frac{1}{\sqrt{\hat{\lambda}}}, \hat{\boldsymbol{\theta}}\right) = \sum_{\mathbf{x}_i \in X \cap A} \frac{1}{\sqrt{\hat{\lambda}(\mathbf{x}_i, X)}} - \int_A \sqrt{\hat{\lambda}(\mathbf{u}, X)} d\mathbf{u}, \tag{11}$$

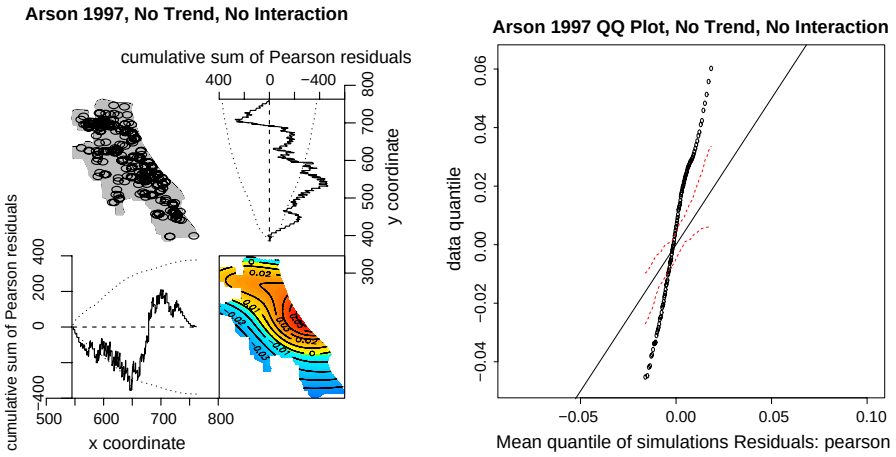

Fig. 12 The 1997 arson fires Pearson residual plots are used to evaluate the trend, and the *QQ*-plot is used to evaluate the interaction. The plots shown here are for a model with no trend and no interaction

where $\hat{\boldsymbol{\theta}}$ is the parameter estimate of a fitted parametric model as in (6). A is, again, a subset of the domain, and X is the set of all locations observed in the point pattern. The $\lambda(\mathbf{u}, X)$ is the Papangelou conditional intensity at the location $\mathbf{u}$ given the outcomes at all other spatial locations and is defined in Baddeley et al. (2005). The residuals have the nice property that they sum to zero if the chosen model is correct. They are not only computed at all locations in the domain but also at "dummy" points since information may be gained not only from where an event occurs but also from where events do not occur. As a result, the computing procedures involving these residuals can be lengthy. While several types of residuals can be computed, we will use the Pearson residuals throughout our analysis.

4.1.1 Evaluating trend

From `spatstat`, a matrix of plots based on the Pearson residuals is used to evaluate the trend. Arson fires in 1997 fitted with constant trend (or homogeneous intensity) and no interaction is used as the baseline comparison in Fig. 12. The top left panel is the *mark plot* that plots a circle proportional to the "size" of the residual at each event location. Circles with an unusually large radius may indicate outliers of the model. The bottom right panel is a *contour plot* of a smoothed residual field that applies a smoothing kernel to the residual measure. It can be interpreted similarly to a topographical map with large contours indicating an area with large residuals. An ideal residual field has most of its contours close to zero.

Finally, the upper right and lower left plots are *lurking variable plots* that help indicate whether or not the presence of a particular variable is needed in the model. The residuals can be plotted against any covariate, but it is common to plot them against the x and y coordinates. The dotted envelopes are 2σ-limits based on the variance of the observed point pattern for an inhomogeneous Poisson point process and are computed using Pearson residuals, so they will be the same for any Poisson interaction model

tested on the data. A cumulative residual function of the residuals against the variable of interest is computed, and if this lies outside of the envelope, this is evidence that the variable should be included in the trend portion of the model. In Fig. 12, it is obvious that both the x and y components of the location should be included in the trend. Note that Baddeley et al. (2005) have stated that envelopes for non-Poisson models are still under construction, so models with an interpoint interaction will be more difficult to assess.

4.1.2 Evaluating interaction

The second plot in Fig. 12 is a *QQ*-plot to diagnose the interaction, or distributional, component of the proposed model. To assess the fit of a distribution, a *QQ*-plot is appropriate. The empirical quantiles of the smoothed residual field are compared to the quantiles expected under the specified interaction computed with Monte Carlo simulations (we use 50 replicates). The dashed lines represent critical intervals for pointwise significance used to detect significant deviation from the model. Note that misspecification of the trend or the interaction can distort the results in either the trend or *QQ*-plots. Thus, in Fig. 12 it is difficult to say whether it is the lack of trend or lack of interaction that is creating such a poor result. The reader should also beware that the procedure to obtain *QQ*-plots is computationally intensive and is not yet able to accommodate more complicated models, such as those including the marks of the locations.

4.2 Results of models tested

Models were tested for arson, accident, and lightning caused fires in the years 1988 and 1997. In all 6 cases, the model with no trend or interaction was a poor fit, similar to the outcome in Fig. 12. To conserve space, only the results of the 1997 arson data will be presented in detail, but the technique used to select the appropriate models in each case was the same. A description of our model selection methodology will be given, and the results for the 1997 arson data will be shown for illustration.

4.2.1 Model selection

Two main classes of models were explored—those with parametric and nonparametric trend. The interaction component was generally assumed to be Poisson since diagnostic plots for non-Poisson interactions are not yet available. For exploration sake, a Geyer interaction to describe clustering of points was tested with the 1997 arson data. Thus, modeling the trend component in order to estimate the intensity accurately in the K and K-cross functions was our primary interest. We first experimented with parametric models for trend. Several classes of functions were investigated, and for each potential trend model, the residual plots were examined. The class of functions with the best fit was the exponential of a polynomial. The exponential of first through sixth degree polynomials were fit to each of the six datasets mentioned above and to the fire locations occurring in each year. The best residual diagnostic plots for every dataset were for the

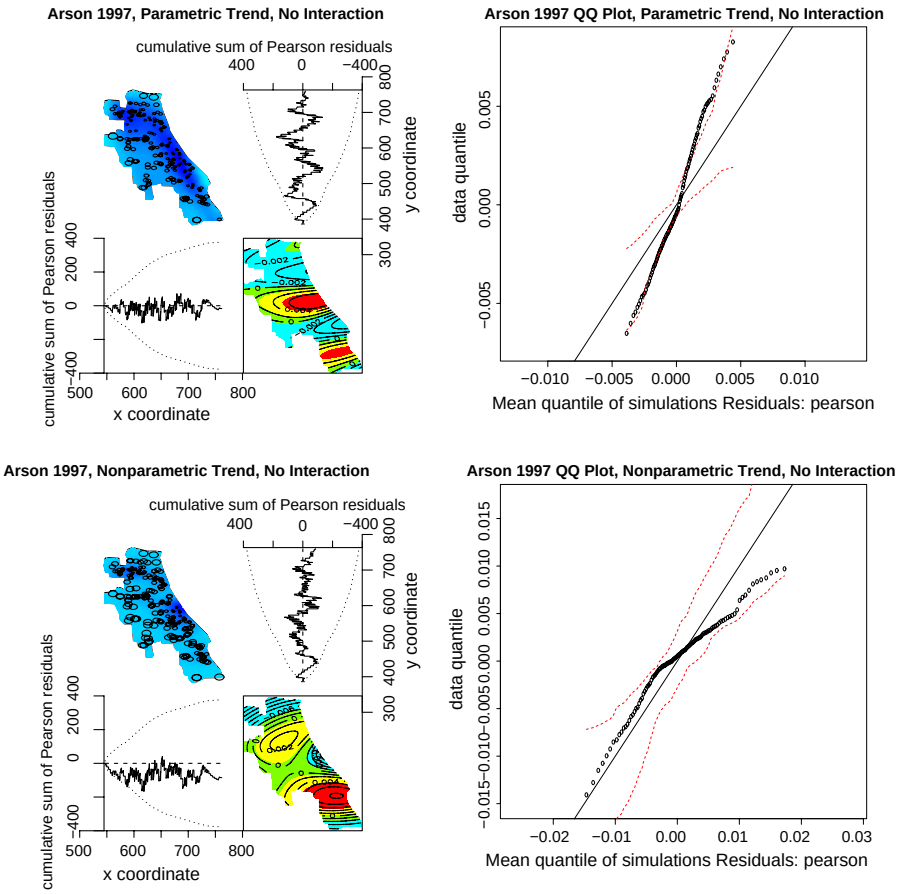

Fig. 13 The 1997 arson fire Pearson residual plots for the parametric and nonparametric ($\sigma = 15$) trend models. Both have no interaction

fourth or fifth degree polynomials with the fifth degree appearing to have slightly better contour plots. The sixth degree polynomial fit deteriorated, possibly due to overfitting. The top panels of Fig. 13 show the exponential of a fifth degree polynomial trend with Poisson interaction residual plots for the 1997 arson data. The improvement over the fit with no trend in Fig. 12 is obvious. This was the trend used in estimating the parametric intensities for the K and K-cross functions.

We also modeled the trend using a nonparametric density estimator with Gaussian kernel. The choice of the bandwidth σ determines the success of the fit. We began our search for an appropriate σ by trying values of σ between 0 and 2 in increments of 0.1. None of these values of σ produced well-behaved residual plots. We extended our search to values of σ up to 100 in increments of 5. We found that very small values of σ and very large values of σ produced poor fits, as mentioned at the end of Sect. 2.2. The best fits resulted from using σ's between 10 and 20, depending on the dataset under consideration. For the arson 1997 data, the σ yielding the best residual plots was

Table 2 Summary of models for each dataset examined in which the residuals for both the lurking variable plots for x and y and the QQ-plots remained within the envelopes

Dataset	Trend component	Interaction
Arson 1988	Nonparametric $\sigma = 15$	Poisson
Arson 1997	Nonparametric $\sigma = 15$	Poisson
Lightning 1988	Parametric	Poisson
Lightning 1997	Nonparametric $\sigma = 20$	Poisson
Accidents 1988	Nonparametric $\sigma = 10$	Poisson
Accidents 1997	Nonparametric $\sigma = 10$	Poisson

Geyer interaction models are not considered when choosing a best model since envelopes in the lurking variable plots are not available

$\sigma = 15$ as shown at the bottom of Fig. 13. There is also noticeable improvement in the QQ-plot compared to the parametric model, but the parametric model has smaller residuals.

For models with a Poisson interaction, Table 2 summarizes which model is best (parametric or nonparametric) for each of the 6 datasets based on the criteria that the residuals remain inside the envelopes in the diagnostic plots. It is interesting to note that the same model works the best in most cases for a given cause for both years. However, more models than just these could be considered. In the future, the residuals software will be capable of handling models that also include information on the marks of each point. Including information on covariate data such as temperature and rainfall and building models based on larger datasets would both be of interest. Finally, a more rigorous selection procedure for the bandwidth σ in the nonparametric trend should be explored.

4.2.2 Non-Poisson interaction component

For comparison sake, we wanted to explore further the impact of changing the interaction component to a distribution that would model the clustering interaction. Both the parametric and nonparametric trends for the 1997 arson data were tested with a Geyer saturation interaction distribution (Geyer 1999). The values that must be specified for the Geyer distribution are r, the interaction radius, and s, the saturation threshold. Both must be positive real numbers. The interaction radius defines the maximum distance that two points can be located from each other and still be considered close neighbors. When s is zero, the Geyer distribution reduces to the Poisson distribution, but when s is a finite positive number, it can lead to a model describing a clustered point process. We chose $s = 2$ as the saturation threshold and experimented with various values of r between 0 and 2. The value of r producing the best diagnostic plots was $r = 0.10$. Methods for estimating r and s from the data are given in Baddeley and Turner (2000). The results are given in Fig. 14.

The trend plots for the parametric and nonparametric trends appear similar, with the cumulative function of the residuals falling sharply in the x lurking variable plot. No envelopes are available here to gauge the degree of departure from zero, but the range on the y axis of both lurking variable plots is smaller than that of the Poisson

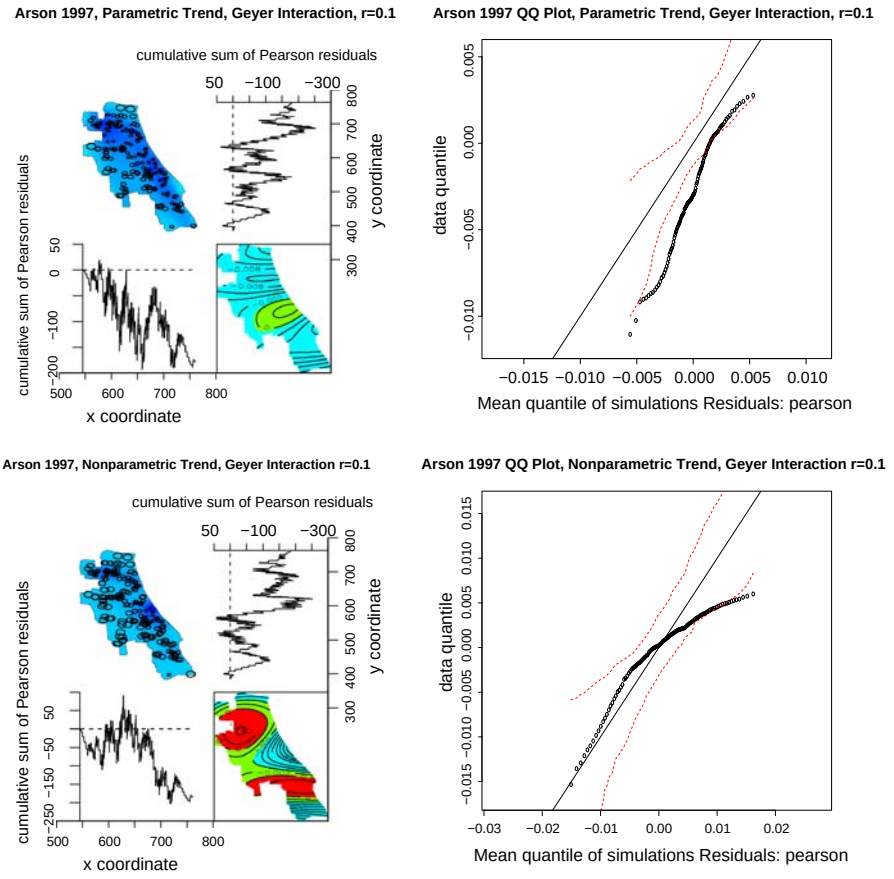

Fig. 14 The 1997 arson fire Pearson residual plots for the parametric and nonparametric ($\sigma = 15$) trend and a Geyer saturation interaction (with parameters $r = 0.10$ and $s = 2$)

interaction models. The *QQ*-plot for the nonparametric Geyer interaction model is significantly better than its parametric counterpart. So, while it is difficult to compare these models to those with Poisson interaction, it does appear that perhaps a different trend may be a better fit with this particular interaction component.

5 Discussion

From comparing the results of the two different analyses of the SJRWMD data and the simulation study, we can see that the homogeneous K-function and K-cross function do not realistically represent the natural phenomenon of clustering. When clustering is present, the K-function analysis with or without the assumption of a homogeneous point process will identify clustering around small distances from a given event. The behavior of the estimated inhomogeneous K-function at large distances indicates that the true K-function for that point process is closer to CSR than that of the homogeneous

estimator. Similarly, the K-cross function overestimates the attraction between two types of points when the intensity is estimated homogeneously.

There is a synonymous relationship between the presence of clustering and the intensity function of a point process. The presence of clustering could be identified at certain places where the intensity is significantly greater than the average intensity across the entire domain. If the intensity of a point process is truly constant throughout the entire domain with no interactions, then there should be no clustering of events. When clustering is truly assumed to be present then to accurately estimate the degree of clustering with a measure such as the K-function or K-cross function one must compute an appropriate intensity function that incorporates these changing densities of events. By accurately estimating the true intensity of the domain, the results of testing departure from CSR are more representative of the true phenomenon of the spatial point process.

An appropriate intensity for a given point process can be chosen via residual analysis. Our analysis has suggested that a trend should be included in a model for wildfire ignitions. The various models we have fitted to the SJRWMD data are rather simple, but they have been shown to capture some of the characteristics of wildfire occurrences. With future improvements in computing techniques, we expect to be able to fit more complex and realistic models.

Acknowledgements The authors are grateful to Marcia Gumpertz, Guest Editor, and two anonymous referees, for very helpful comments that improved the manuscript. The authors thank the Southern Research Station of the USDA Forest Service for making the data available and for partially funding this research by SRS grant 01-CA-11330143-412. This research was also partially supported by NSF grant DMS-0504896.

References

Agee JK (2003) Monitoring postfire tree mortality in mixed-conifer forests of Crater Lake, Oregon. Nat Areas J 23:114–120

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

Baddeley A, Turner R (2000) Practical maximum pseudolikelihood for spatial point patterns (with discussion). Aust N Z J Stat 42:283–322

Baddeley A, Turner R (2005) Spatstat: an R package for analyzing spatial point patterns. J Stat Softw 12:1–42

Baddeley A, Turner R, Møller J, Hazelton M (2005) Residual analysis for spatial point processes (with discussion). J Roy Stat Soc Ser B 67:617–666

Brose P, Wade DD (2002) Potential fire behavior in pine flatwood forests following three different fuel reduction techniques. For Ecol Manage 163:71–84

Butry DT, Gumpertz ML, Genton MG (2008) The production of large and small wildfires. Chapter 5. In: Holmes TP, Prestemon JP, Abt, KL (eds) Economics of forest disturbances: wildfires, storms, and invasive species. Springer: Dordrecht, The Netherlands

Diggle PJ (1983) Statistical analysis of spatial point patterns. Academic Press, London

Diggle PJ (1985) A kernel method for smoothing point process data. Appl Stat 34:138–147

Genton MG, Butry DT, Gumpertz ML, Prestemon JP (2006) Spatio-temporal analysis of wildfire ignitions in the St. Johns River Water Management District, Florida. Int J Wildland Fire 15:87–97

Geyer CJ (1999) Likelihood inference for spatial point processes. Chapter 3. In: Barndorff-Nielsen OE, Kendall WS, Van Lieshout MNM (eds) Stochastic geometry: likelihood and computation. Chapman and Hall/CRC, Monographs on Statistics and Applied Probability, 80:79–140

Haines TK, Busby RL, Cleaves DA (2001) Prescribed burning in the south: trends, purpose, and barriers. South J Appl Forest 25:149–153

Møller J, Waagepetersen R (2003) Statistical inference and simulation for spatial point processes. Chapman and Hall, Boca Raton
Outcalt KW, Wade DD (2004) Fuels management reduces tree mortality from wildfires in southeastern United States. South J Appl Forest 28:28–34
Podur J, Martell DL, Csillag F (2003) Spatial patterns of lightning-caused forest fires in Ontario, 1976–1998. Ecol Modell 164:1–20
Prestemon JP, Pye JM, Butry DT, Holmes TP, Mercer DE (2002) Understanding wildfire risks in a human-dominated landscape. For Sci 48:685–693
Ripley BD (1977) Modelling spatial patterns (with discussion). J Roy Stat Soc Ser B 39:172–212
Ripley BD (1981) Spatial statistics. Wiley, New York
Schabenberger O, Gotway CA (2005) Statistical methods for spatial data analysis. Chapman and Hall, Boca Raton

Author Biographies

Amanda S. Hering is a Ph.D. student working under the supervision of Prof. Marc G. Genton in the Department of Statistics at Texas A&M University. Her research interests include circular statistics and space-time modeling of environmental data, particularly with non-Gaussian distributions. She received a B.S. in Mathematics with a minor in Environmental Studies at Baylor University, Waco, Texas, in 1999 and an M.S. in Statistics at Montana State University, Bozeman, in 2002.

Cynthia L. Bell is a statistician at the University of Texas Health Science Center at Houston. She is currently researching indicators and treatments for pediatric hypertension. Her interests include applications of spatial statistics to ecology and epidemiology, econometrics and small-sample methods. She received her M.S. in Statistics from Texas A&M University in 2006.

Marc G. Genton is Professor of Statistics in the Department of Econometrics at the University of Geneva and Associate Professor in the Department of Statistics at Texas A&M University. His research interests include spatio-temporal statistics, robustness, multivariate analysis, skewed multivariate distributions, and data mining. In Fall 2006, he organized a workshop on Multivariate Methods in Environmetrics. He earned his Ph.D. in Statistics from the Swiss Federal Institute of Technology at Lausanne in 1996.