Spatio-temporal point process analysis of Mexico State wildfires

Luis Ramón Munive-Hernández*[1] and Antonio Villanueva-Morales†[2]

[1] Departamento de Matemáticas, Universidad Autónoma Metropolitana

[1] Colegio de Ciencias y Humanidades, Universidad Autónoma de la Ciudad de México

[2] Departamento de Estadística, Matemática y Cómputo, Universidad Autónoma Chapingo

Abstract

Wildfires are an example of a phenomenon that can be investigated using point process theory. We analyze public data from the National Forestry Commission. It consists of wildfire records, specifically their coordinates and dates of occurrence in Mexico State from 2010 to 2018. The spatial component was examined and we found that wildfires tend to cluster. Afterwards, a time series analysis was conducted. This shows that the data comes from a stationary stochastic process. Finally, some spatio-temporal features that demonstrate the point process' regular behaviour in space and time were investigated. This research could be a reference to describe wildfire behaviour in a specific space and time.

Keywords— Environmental statistics, point processes, spatio-temporal statistics, wildfires.

1 Introduction

Wildfires are complex phenomena with serious socio-environmental consequences, including economic and biodiversity losses, among others. Anthropogenic factors are responsible for nearly all wildfires in Mexico State, according to data from the National Forestry Commission (Conafor, its Spanish acronym) [1] (see Figure 1).

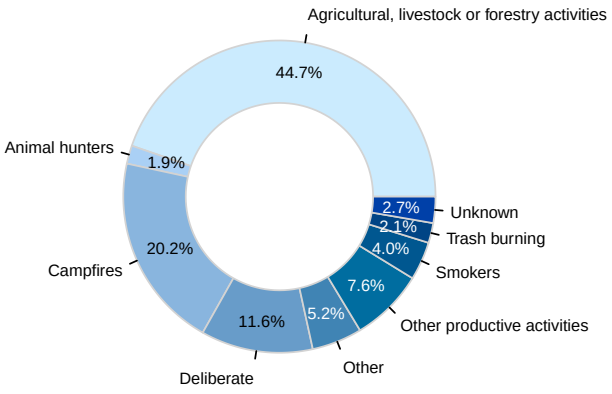


Figure 1: Mexico State wildfire causes (2010-2018).

There is plenty of specialized literature available on wildfires (see [2] and [3]). The authors of [4] use a logistic regression model to assess the risk of wildfire in Puebla, Mexico, taking into account land cover, meteorological, topographic, and social variables. Using two different data sources: Conafor's open data and Modis' (Moderate Resolution Imaging Spectroradiometer) data, the authors of [5] show that wildfire spatial patterns in Mexico tend to cluster. The spatial and temporal relationships between Conafor's wildfire records from 2005 to 2015 and the Standardized Precipitation-Evapotranspiration Index (SPEI) were investigated [6]. Machine learning techniques were used to determine the wildfire propensity in Mexico using Conafor's open data [7].

The spatio-temporal behaviour of wildfires could be critical for improving fire management strategies. The point processes approach can be used to model random events in time, space, or space-time, such as wildfires. In this study, we used point processes theory to describe the spatio-temporal behaviour of wildfires in Mexico State from 2010 to 2018.

*cbi2202800068@izt.uam.mx, luis.ramon.munive@alumnos.uacm.edu.mx

†avillanuevam@chapingo.mx

2 Point processes basic theory

A point process is a random set in which the number of points and their locations are both random [8, p. 12]. A point process could occur in any completely separable metric space $\mathcal{S}$, such as d-dimensional Euclidean space $\mathbb{R}^d$.

Definition 1 (Point process)**.** The point process Y, with state space $\mathcal{S}$, is a measurable mapping from a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ to the measure space of the point process' realizations equipped with the counting measure, $\left(\mathcal{Y}_{\mathcal{S}}^{\#}, \mathscr{B}\left(\mathcal{Y}_{\mathcal{S}}^{\#}\right), \mu_{\#}\right)$. Where $\mathcal{Y}_{\mathcal{S}}^{\#} = \{\mu_{\#} : \mathscr{B}(\mathcal{S}) \to \mathbb{N} \mid \ \mu_{\#}(A) < \infty, A \in \mathscr{B}(\mathcal{S})\}$ is the space of all finite counting measures on $\mathscr{B}(\mathcal{S})$.

The commutative diagram in Figure 2 illustrates the point process definition.

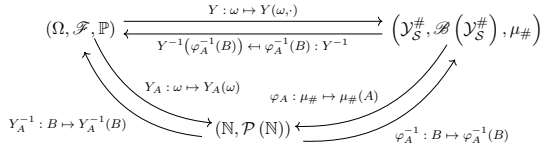


Figure 2: Commutative diagram of point process definition.

The mapping φ_A takes measures $\mu_{\#} \in \mathcal{Y}_{\mathcal{S}}^{\#}$ and maps them into $\mu_{\#}(A)$. As a result, the mapping φ_A in terms of the point process Y is $\varphi_A : Y(\omega, \cdot) \mapsto Y(\omega, A)$.

Furthermore, the commutative diagram reveals the equivalences: $Y(\omega, A) = \varphi_A\left(Y(\omega, \cdot)\right) = Y_A(\omega)$ and $Y_A^{-1}(B) = Y^{-1}\left(\varphi_A^{-1}(B)\right)$, for any $B \in \mathcal{P}(\mathbb{N})$, [9, pp. 8–9], [10, p. 13]. The following are some fundamental properties of a point process [10, pp. 7–8]:

i. Is additive, this is

$$Y(\omega, A_1 \cup A_2) = Y(\omega, A_1) + Y(\omega, A_2),$$

whenever $A_1 \cap A_2 = \varnothing$, $A_1, A_2 \subset \mathcal{S}$ and of course

$$Y(\omega, \varnothing) = 0.$$

ii. Is locally finite

$$\mathbb{P}\left(Y(\omega, A) < \infty\right) = 1,$$

for any $A \subset \mathcal{S}$.

iii. Is simple

$$\mathbb{P}\left(Y(\omega, \{\boldsymbol{s}\}) \leq 1\right) = 1,$$

for any point $\boldsymbol{s} \in \mathcal{S}$.

For simplification, we will write $Y(\omega, A) = Y(A)$ in the foregoing. When the point process Y is observed, we have a point pattern denoted by $\boldsymbol{Y}$.

In order to generate models, some assumptions about a point process must be made. Stationarity and isotropy are the most important assumptions. The former refers to statistical invariance under translations, whereas the latter refers to statistical invariance under rotations [10, p. 16], [11, pp. 146–147]. Nonetheless, some research on non-stationary and anisotropic processes has been conducted (see [12] and [13, ch. 5]).

Definition 2 (Stationary point process)**.** A point process Y on $\mathcal{S}$ is stationary if, for any fixed $\boldsymbol{s} \in \mathcal{S}$, the distribution of the process $Y + \boldsymbol{s}$ is identical to the distribution of Y.

2.1 Poisson process

The general Poisson point process in some space $\mathcal{S}$ can be defined as follows [10, p. 12], [11, pp. 300–301].

Definition 3 (Poisson process)**.** The Poisson process Y on $\mathcal{S}$ with intensity measure Λ is a point process such that:

i. For every compact set $A \subset \mathcal{S}$, the random variable $Y(A) \sim \text{Poisson}\left(\Lambda(A)\right)$.

ii. If $A_1, \ldots, A_n \subset \mathcal{S}$ are disjoint compact sets, then $Y(A_1), \ldots, Y(A_n)$ are independent random variables.

Where the intensity measure Λ is defined, for any $A \subset \mathcal{S}$, as $\Lambda(A) = \mathbb{E}\left(Y(A)\right)$.

If the state space is $\mathcal{S} = \mathbb{R}^2 \times \mathbb{R}_+$ and the expected value of the point process Y in $S \times T$, with $S \subset \mathbb{R}^2$ and $T \subset \mathbb{R}_+$, can be written as follows:

$$\mathbb{E}\left(Y(S \times T)\right) = \lambda\, \mu_L(S)\, \mu_L(T),$$

where $\lambda > 0$ and μ_L is the Lebesgue measure, then we have the spatio-temporal homogeneous Poisson point process [14, pp. 9–10].

The simplest stochastic mechanism for generating point patterns is the homogeneous Poisson point process. As a data model, it is almost never plausible. Regardless, it is the fundamental reference or benchmark model of a point process [8, p. 53].

The homogeneous Poisson point process is also known as complete spatial (or spatio-temporal) randomness. Additionally, the Poisson point process is stationary and isotropic [10, p. 16].

Figure 3 depicts a spatial point pattern generated by a homogeneous Poisson point process.

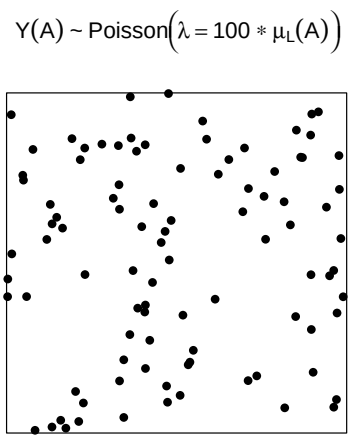


Figure 3: Simulation of a spatial homogeneous Poisson process.

3 Point pattern's data analysis

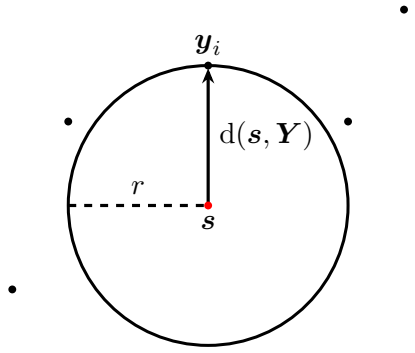
Distances between points are a straightforward way to examine a point pattern. The most common statistics used in exploratory analysis of a point pattern are as follows.

3.1 Empty-space function F

Let Y be a stationary point process on $\mathcal{S}$. The shortest distance between a given point $\boldsymbol{s} \in \mathcal{S}$ and the nearest observed point $\boldsymbol{y}_i \in \boldsymbol{Y}$ is denoted as $\mathrm{d}(\boldsymbol{s}, \boldsymbol{Y}) = \min_i \{\|\boldsymbol{s} - \boldsymbol{y}_i\|\}$. It is called the empty-space distance, spherical contact distance, or simply contact distance [8, p. 83], [10, pp. 21–22], [11, pp. 261–262]. Note that

$$\mathrm{d}(\boldsymbol{s}, \boldsymbol{Y}) \leq r \Leftrightarrow Y\left(B_r(\boldsymbol{s})\right) > 0, \tag{1}$$

where $B_r(\boldsymbol{s})$ is the neighborhood of radius r centered on $\boldsymbol{s}$.



$B_r(\boldsymbol{s})$

Figure 4: Empty-space distance illustration.

In other words, as shown in Figure 4, the empty-space distance satisfies the logical equivalence of the biconditional 1, $\mathrm{d}(\boldsymbol{s}, \boldsymbol{Y}) > r \Leftrightarrow Y\left(B_r(\boldsymbol{s})\right) = 0$.

Moreover, because $\{Y\left(B_r(\boldsymbol{s})\right) > 0\}$ is measurable, the event $\{\mathrm{d}(\boldsymbol{s}, \boldsymbol{Y}) \leq r\}$ is measurable, implying that the contact distance is a well-defined random element.

Definition 4 (Empty-space function F)**.** Let Y be a stationary point process on $\mathcal{S}$. The empty-space function F is the cumulative distribution function of the empty-space distance

$$F(r) = \mathbb{P}\left(\mathrm{d}(\boldsymbol{s}, \boldsymbol{Y}) \leq r\right).$$

If Y is a homogeneous Poisson process on $\mathbb{R}^d$ with intensity λ, then the empty-space function is

$$F(r) = 1 - \exp\left(-\lambda\, \mu_L\left(B_1(\boldsymbol{0})\right)\, r^d\right),$$

where $r \geq 0$, $\mu_L\left(B_1(\boldsymbol{0})\right) = \frac{\pi^{d/2}}{\Gamma\left(\frac{d}{2}+1\right)}$ denotes the volume of the unitary d-ball in $\mathbb{R}^d$ and Γ is the usual gamma function.

3.2 Nearest-neighbour function G

The nearest-neighbour distance, denoted by $\mathrm{d}_i = \min_{i \neq j} \left\{\|\boldsymbol{y}_i - \boldsymbol{y}_j\|\right\}$, is the distance between each point $\boldsymbol{y}_i \in \boldsymbol{Y}$ and its nearest neighbour in the set $\boldsymbol{Y} \backslash \{\boldsymbol{y}_i\}$, [8, p. 90], [10, pp. 51–52]. It is worth noting that d_i can also be written as $\mathrm{d}_i = \mathrm{d}\left(\boldsymbol{y}_i, \boldsymbol{Y} \backslash \{\boldsymbol{y}_i\}\right)$, [11, p. 262]. This distance is depicted in Figure 5.

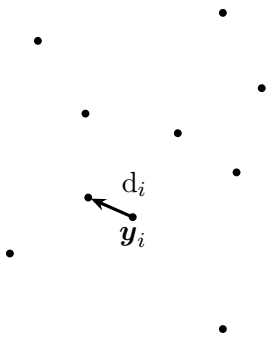


Figure 5: Nearest-neighbour distance illustration.

Definition 5 (Nearest-neighbour function G)**.** Let Y be a stationary point process on $\mathcal{S}$. The nearest-neighbour function G is the cumulative distribution function of the nearest-neighbour distance

$$G(r) = \mathbb{P}\left(\mathrm{d}\left(\boldsymbol{s}, \boldsymbol{Y} \backslash \{\boldsymbol{s}\}\right) \leq r \mid \ \boldsymbol{s} \in \boldsymbol{Y}\right),$$

where $r \geq 0$ and $\boldsymbol{s}$ is any location in the state space $\mathcal{S}$.

If Y is a homogeneous Poisson process on $\mathbb{R}^d$ with intensity λ, then the nearest-neighbour function is

$$G(r) = 1 - \exp\left(-\lambda\, \mu_L\left(B_1(\mathbf{0})\right)\, r^d\right).$$

In this case, we have that $F(r) = G(r)$, i.e., under complete spatial randomness, the points of the Poisson process are independent of each other, so conditioning does not affect them. Therefore, F is equivalent to G, [8, p. 91].

3.3 Intensity

The intensity function describes the first-order properties of a point process [15, p. 623], [16, p. 57].

The average number of points per spatial (or spatio-temporal) unit defines the intensity of a point process. In this regard, intensity is analogous to the expected value of a random variable [10, p. 26].

Similarly, we can investigate the analogue of a point process' variance or covariance throughout the second-order properties.

As we will see in the following, the intensity measure Λ of a point process Y is clearly a set function, whereas the "instantaneous" intensity function λ is an atomic function.

Definition 6 (First-order intensity)**.** Let Y be a point process on $\mathcal{S}$. The first-order intensity is defined as

$$\lambda(\boldsymbol{s}) = \lim_{\nu(\mathrm{d}\boldsymbol{s})\to 0} \frac{\mathbb{E}\left(Y(\mathrm{d}\boldsymbol{s})\right)}{\nu\left(\mathrm{d}\boldsymbol{s}\right)},$$

where ν is a suitable measure on $(\mathcal{S}, \mathscr{B}(\mathcal{S}))$ and $\mathrm{d}\boldsymbol{s}$ defines a infinitesimally small region around $\boldsymbol{s}$.

If Y is a point process on $\mathbb{R}^d$ with intensity measure Λ, it satisfies

$$\Lambda(A) = \int_A \lambda(\boldsymbol{s})\, \mu_L\left(\mathrm{d}\boldsymbol{s}\right),$$

for some function λ and any $A \subset \mathbb{R}^d$. Then λ is called the intensity function of Y [10, p. 27]. If λ is constant, then Y is said to be homogeneous, otherwise is said to be inhomogeneous [17, p. 40]. Likewise, if the intensity function exists, we can interpret it as follows:

$$\mathbb{P}\left(Y(\mathrm{d}\boldsymbol{s}) > 0\right) \approx \mathbb{E}\left(Y(\mathrm{d}\boldsymbol{s})\right) \approx \lambda(\boldsymbol{s})\, \mu_L\left(\mathrm{d}\boldsymbol{s}\right).$$

The K function and pair correlation are both second-moment properties, so the second-order intensity must be defined [16, p. 57].

Definition 7 (Second-order intensity)**.** Let Y be a point process on $\mathcal{S}$. The second-order intensity is defined as

$$\lambda_2(\boldsymbol{s}, \boldsymbol{u}) = \lim_{\substack{\nu(\mathrm{d}\boldsymbol{s})\to 0 \\ \nu(\mathrm{d}\boldsymbol{u})\to 0}} \frac{\mathbb{E}\left(Y(\mathrm{d}\boldsymbol{s})\, Y(\mathrm{d}\boldsymbol{u})\right)}{\nu\left(\mathrm{d}\boldsymbol{s}\right)\, \nu\left(\mathrm{d}\boldsymbol{u}\right)}.$$

We already have the fundamental elements for defining the following pair of second-order properties.

3.4 K function

The K function counts the number of locations within a certain radius of a given point (see Figure 6), [11, p. 226], [18, p. 171]. Ripley defined it in [19]. We present the following definition [8, p. 92], [16, pp. 57–58].

Definition 8 (K function)**.** Let Y be a stationary and isotropic point process on $\mathcal{S}$ with intensity λ. The K function is defined as

$$K(r) = \frac{1}{\lambda}\mathbb{E}\left(Y\left(\boldsymbol{Y} \cap B_r(\boldsymbol{s})\backslash\{\boldsymbol{s}\}\right) \mid \ \boldsymbol{s} \in \boldsymbol{Y}\right),$$

where $r \geq 0$ and $\boldsymbol{s}$ is any location in $\mathcal{S}$.

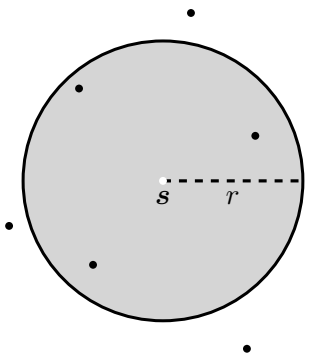


Figure 6: K function illustration.

If $\mathcal{S} = \mathbb{R}^d$ and the point process Y is assumed to be stationary, then hold $\lambda_2\left(\boldsymbol{s}, \boldsymbol{u}\right) = \lambda_2\left(\boldsymbol{s} - \boldsymbol{u}\right)$. Also, if Y is isotropic, hence $\lambda_2(\boldsymbol{s} - \boldsymbol{u}) = \lambda_2(r)$, where $r = \|\boldsymbol{s} - \boldsymbol{u}\|$. These conditions implies that [15, p. 633], [16, p. 58],

$$\lambda\, K(r) = \frac{d\, \mu_L\left(B_1(\mathbf{0})\right)}{\lambda} \int_0^r \lambda_2(z)\, z^{d-1}\, \mathrm{d}z. \qquad (2)$$

The above expression provides a relationship between the K function and the second-order intensity under the assumptions of stationarity and isotropy.

If Y is a homogeneous Poisson process on $\mathbb{R}^d$, then the K function is [10, p. 38],

$$K(r) = \mu_L\left(B_1(\mathbf{0})\right)\, r^d.$$

3.5 Pair correlation function g

In general, the pair correlation function is a quotient of probabilities; that is, the probability of observing a pair of points separated by a given distance is divided by the same probability, assuming a Poisson point process [8, p. 94]. In the strictest sense, it is neither a distribution nor a correlation function [16, p. 57].

Some authors consider the pair correlation function to be the most informative second-order property because it provides information more simply than, say, the K function [20, p. 218]. We present the following definition [10, pp. 33–34], [17, p. 41].

Definition 9 (Pair correlation function g)**.** Let Y be a point process on $\mathcal{S}$ with intensity function λ and second-moment density g_2. The pair correlation function g is defined as

$$g(\boldsymbol{s}, \boldsymbol{u}) = \frac{g_2(\boldsymbol{s}, \boldsymbol{u})}{\lambda(\boldsymbol{s})\, \lambda(\boldsymbol{u})},$$

for any $\boldsymbol{s}, \boldsymbol{u} \in \boldsymbol{Y}$, where the second-moment density is such that

$$\nu_{[2]}(C) = \int_C g_2(\boldsymbol{s}, \boldsymbol{u})\, \nu\,(\mathrm{d}\boldsymbol{s})\, \nu\,(\mathrm{d}\boldsymbol{u})\,,$$

for any compact set $C \subset \mathcal{S} \times \mathcal{S}$, where ν is a suitable measure on $(\mathcal{S}, \mathscr{B}(\mathcal{S}))$ (e.g., if $\mathcal{S} = \mathbb{R}^d$, so $\nu = \mu_L$), and $\nu_{[2]}(A_1 \times A_2) = \mathbb{E}\,(Y(A_1)\, Y(A_2)) - \mathbb{E}\,(Y\,(A_1 \cap A_2))$, with $A_1, A_2 \subset \mathcal{S}$, is the second factorial moment measure of Y.

If Y is stationary and isotropic, it follows from 2 that [16, p. 58], [20, p. 219],

$$g(r) = \frac{K'(r)}{\mu_L\,(B_1(\mathbf{0}))\, d\, r^{d-1}}.$$

We can define g graphically by taking two concentric circles with radius r and $r + \Delta r$, where Δr is a small increment, and counting the points that fall within the ring (see Figure 7), [11, pp. 225–226].

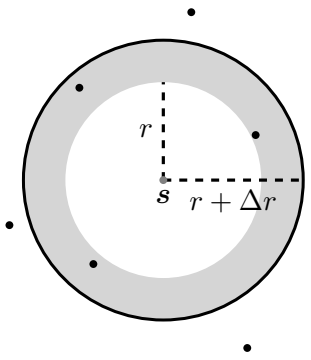


Figure 7: Pair correlation function g illustration.

If Y is stationary and isotropic, the expected number of locations in the ring is $\lambda\, K(r + \Delta r) - \lambda\, K(r)$. Dividing it by the expected value of points assuming a Poisson process, we obtain

$$\begin{aligned} g_{\Delta r}(r) &= \frac{\lambda\,(K(r + \Delta r) - K(r))}{\lambda\, \mu_L\,(B_1(\mathbf{0}))\left((r + \Delta r)^d - r^d\right)} \\ &= \frac{K(r + \Delta r) - K(r)}{\mu_L\,(B_1(\mathbf{0}))\left(\sum_{k=0}^{d} \binom{d}{k} r^{d-k}\,(\Delta r)^k - r^d\right)}. \end{aligned} \tag{3}$$

All binomial expansion components in the denominator of the second line in 3 lose significance except for $d\, r^{d-1}\Delta r$, so

$$g_{\Delta r}(r) \approx \frac{K(r + \Delta r) - K(r)}{\mu_L\,(B_1(\mathbf{0}))\, d\, r^{d-1}\Delta r}.$$

Taking the following limit, we get

$$\begin{aligned} \lim_{\Delta r \to 0} g_{\Delta r}(r) &\approx \lim_{\Delta r \to 0} \frac{K(r + \Delta r) - K(r)}{\mu_L\,(B_1(\mathbf{0}))\, d\, r^{d-1}\, \Delta r} \\ &= \frac{K'(r)}{\mu_L\,(B_1(\mathbf{0}))\, d\, r^{d-1}} \\ &= g(r). \end{aligned}$$

If Y is a homogeneous Poisson process on $\mathbb{R}^d$, then the pair correlation function is $g(r) = 1$.

4 Wildfires' data analysis

Conafor data are licensed for free use (see details in https://datos.gob.mx/libreusomx). It includes wildfire geographical coordinates and dates, as well as variables like forest type affected and severity, among other things.

4.1 Spatial analysis

This spatial analysis focuses on the F and G functions to determine whether the wildfire spatial point pattern is aggregated, complete spatial random, or regular. In addition, the intensity was estimated to support the evidence about point pattern behavior.

Plotting the spatial point pattern is a good starting point for understanding its behavior. Figure 8 shows the spatial point pattern. The wildfires do not appear to be the result of a Poisson process.

There are multiple ways to prove if a point pattern comes from a Poisson point process (see [11, ch. 10]).

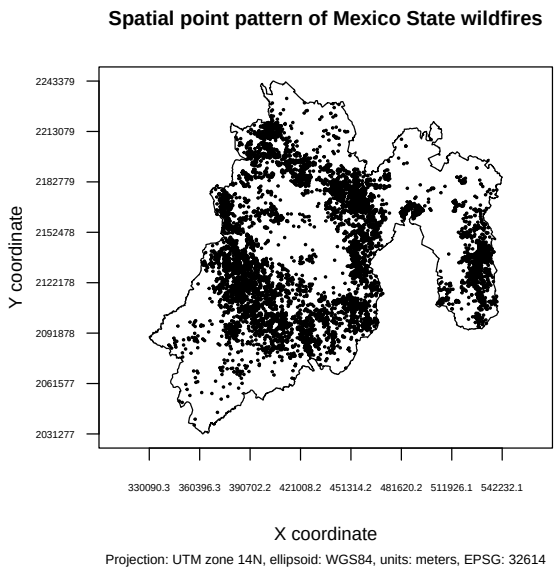


Figure 8: Spatial point pattern of Mexico State wildfires.

The simulation envelopes provide a formal way to decide if the spatial pattern comes from the Poisson process. It is equivalent to performing a hypothesis test. The simulation envelopes are obtained under the assumption of a Poisson process [8, pp. 98–99], [11, pp. 268–271], [18, pp. 161–163].

If the empirical curve falls within the envelope, we can conclude that the point pattern comes from a Poisson process.

Figures 9 and 10 show the estimated F and G functions, as well as the theoretical functions for the Poisson process and simulation envelopes. For this, we use the `R` package `spatstat` [21].

Clearly, the spatial point pattern does not follow the Poisson model.

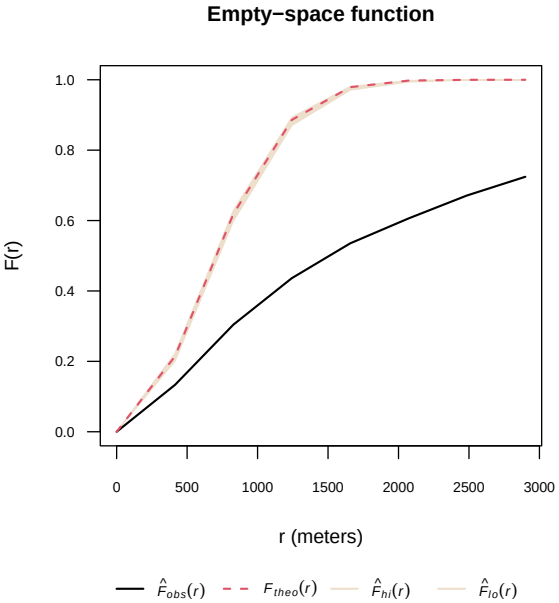


Figure 9: Estimated F function and simulation envelopes.

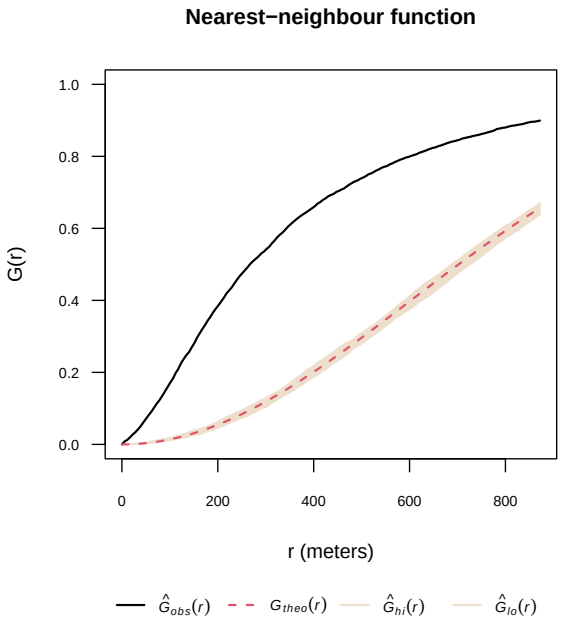


Figure 10: Estimated G function and simulation envelopes.

In Figure 9 note that $\widehat{F}_{\text{obs}}(r) < F_{\text{theo}}(r)$, i.e., the point pattern has longer empty-space distances than a Poisson process. This suggests a clustered point pattern [8, p. 86]. While in Figure 10 we observe that $\widehat{G}_{\text{obs}}(r) > G_{\text{theo}}(r)$, i.e., the point pattern has shorter nearest-neighbour distances than a Poisson model, indicating a clustered pattern [8, p. 91].

Figure 11 depicts the estimated intensity using a Gaussian kernel with bandwidth of 17 km. It can be used to locate wildfire hotspots.

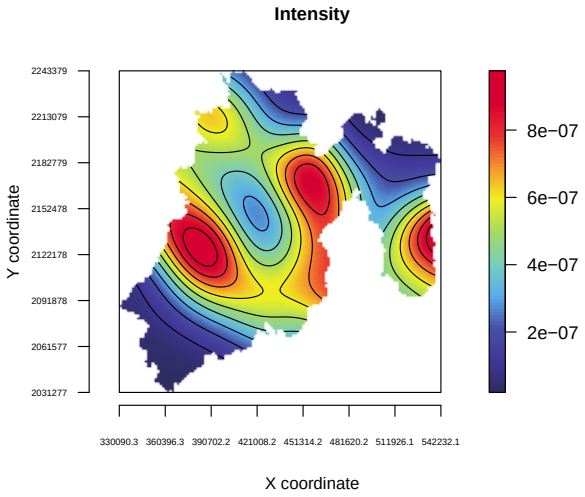


Figure 11: Estimated intensity.

4.2 Time series analysis

This time series analysis was carried out to describe the temporal behavior of wildfires. Figure 12 displays the daily number of wildfires. This immediately suggests that the wildfire time series is seasonal.

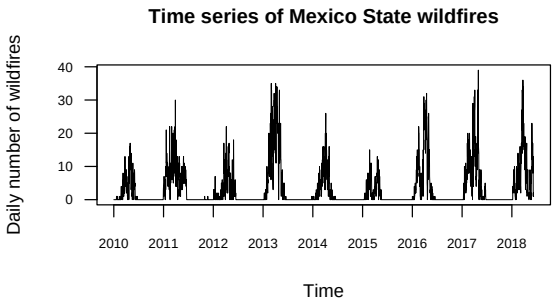


Figure 12: Time series of Mexico State wildfires.

The augmented Dickey-Fuller test is used to prove that the time series is seasonal (see details in [22, pp. 169–173]). This test is included in the `R` package `tseries` [23], where the null hypothesis is that the time series is non-stationary, against the alternative hypothesis that the time series is stationary.

Table 1 displays the results of the augmented Dickey-Fuller test for the wildfire time series, with a significance level of $\alpha = 0.05$.

Test statistic	p-value
-5.1037	< 0.01

Table 1: Augmented Dickey-Fuller test results.

4.3 Spatio-temporal analysis

To demonstrate clustering or regularity in a spatio-temporal point pattern, the space-time inhomogeneous K function (STIK) and space-time pair correlation function (STPC) can be used [14, p. 6], [24]. We employ the `R` package `stpp` [14] for this purpose.

On the assumption that the point process Y on $\mathbb{R}^d$ is second-order stationary, that is, their first-order and second-order properties are invariant under translations, the K function is [24, p. 45],

$$K(r) = d\,\mu_L\left(B_1(\mathbf{0})\right)\int_0^r g(z)\,z^{d-1}\mathrm{d}z. \qquad (4)$$

In addition, a spatio-temporal point process is second-order intensity reweighted stationary and isotropic if its intensity function is bounded away from zero, and its g function is solely determined by (u, v), where $u = \|\boldsymbol{s}_i - \boldsymbol{s}_j\|$ and $v = |t_i - t_j|$, with $\boldsymbol{s}_i, \boldsymbol{s}_j \in \mathbb{R}^2$, $t_i, t_j \in \mathbb{R}_+$, [14, p. 3].

Let Y be a second-order intensity reweighted stationary and isotropic spatio-temporal point process with intensity λ; then, from 4, its STIK function is, [14, p. 6], [24, p. 45],

$$K_{ST}(u, v) = 2\pi\int_0^v\int_0^u g(w, z)\,w\,\mathrm{d}w\,\mathrm{d}z,$$

where $g(u, v) = \frac{\lambda_2(u,v)}{\lambda(\boldsymbol{s}_i,t_i)\,\lambda(\boldsymbol{s}_j,t_j)}$ is the spatio-temporal pair correlation function g of Y.

For any inhomogeneous spatio-temporal Poisson process with intensity bounded away from zero,

$$K_{ST}(u, v) = \pi u^2 v.$$

Figures 13 and 14 show the estimated STIK function in contour and perspective plots, respectively.

The values $\widehat{K}_{ST}(u, v) - \pi u^2 v$ were plotted in order to use them as a measure of spatiotemporal aggregation or regularity. According to [24, p. 45], $\widehat{K}_{ST}(u, v) - \pi u^2 v < 0$ indicates regularity.

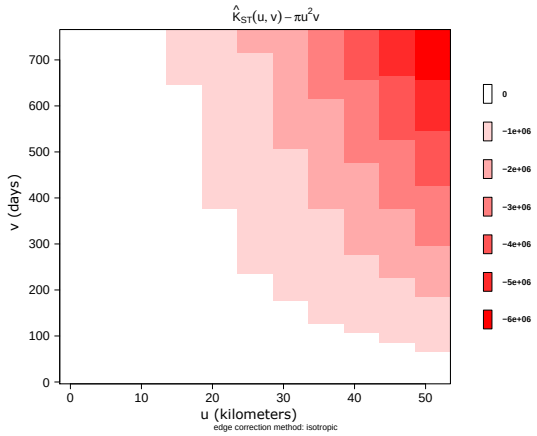


Figure 13: Estimated STIK function contour plot.

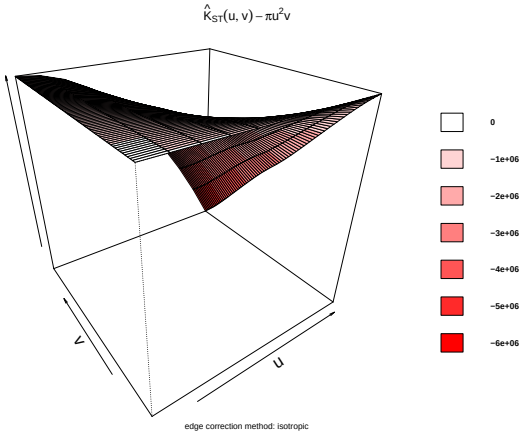


Figure 14: Estimated STIK function perspective plot.

Figures 15 and 16 illustrate estimated STPC function in contour and perspective plots, respectively.

For a spatio-temporal Poisson point process, $g(u, v) = 1$. This reference can be used to determine how much more or less likely it is that a pair of events will occur at specific locations than in a Poisson process of equal intensity [14, p. 3].

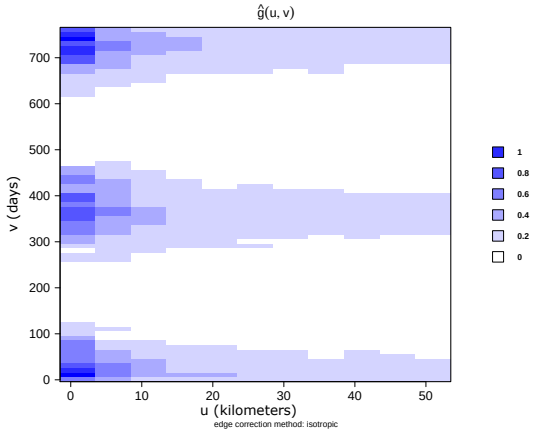


Figure 15: Estimated STPC function contour plot.

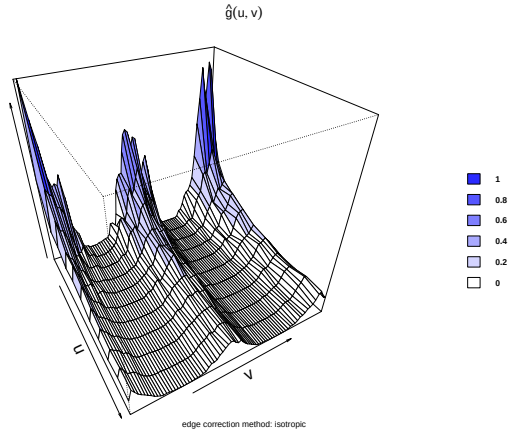


Figure 16: Estimated STPC function perspective plot.

Surface behavior is regular; that is, there is yearly seasonality at distances less than 10 km, implying spatio-temporal regularity.

5 Conclusions and perspectives

The spatio-temporal point pattern of Mexico State wildfires from 2010 to 2018 tends to cluster spatially, as shown by Figures 8, 9, 10, and 11.

While the temporal behavior is stationary, as illustrated in Figure 12 and Table 1, there is a yearly wildfire season during the first semester of each year.

Finally, as shown in Figures 13, 14, 15, and 16, we demonstrate that the spatio-temporal behavior is regular. This means that wildfires tend to occur in the same season and in the same areas each year. This regular spatio-temporal behavior suggests that the underlying point process is predictable in some ways.

This research could be expanded by looking into models such as spatio-temporal log-Gaussian Cox processes [25], which can be used to make spatio-temporal predictions.

Acknowledgments

The authors would like to express their gratitude to the Universidad Autónoma Chapingo.

Appendix

This analysis was performed using the statistical programming language `R` [26]. The developed code is available in the repository:

`https://github.com/LuisMunive/Spatio-temporal-point-process-analysis-of-Mexico-State-wildfires`.

References

[1] Conafor, "Historical yearly series of wildfires 2010-2018 period." Extracted from: https://datos.gob.mx/busca/dataset/incendios-forestales/resource/5720e224-3d0c-4eed-ac65-ea7aac7d72e8, 2018. Date: 2019-08-04.

[2] D. A. Rodríguez-Trejo, "Incendios de vegetación: su ecología, manejo e historia vol. 1," tech. rep., 2014.

[3] D. A. Rodríguez-Trejo, "Incendios de vegetación: su ecología, manejo e historia vol. 2," tech. rep., 2015.

[4] R. L. Carrillo-García, D. A. Rodríguez-Trejo, H. Tchikoué, A. I. Monterroso-Rivas, and J. Santillan-Pérez, "Análisis espacial de peligro de incendios forestales en Puebla, México," *Interciencia*, vol. 37, no. 9, pp. 678–683, 2012.

[5] D. Cisneros-González, G. Pérez-Verdín, M. Pompa-García, D. A. Rodríguez-Trejo, and J. M. Zúñiga-Vásquez, "Spatial modeling of forest fires in Mexico: an integration of two data sources," *Bosque*, vol. 38, no. 3, pp. 563–574, 2017.

[6] M. Pompa-García, J. Camarero J., D. A. Rodríguez-Trejo, and D. J. Vega-Nieva, "Drought and spatiotemporal variability of forest fires across Mexico," *Chinese Geographical Science*, vol. 28, no. 1, pp. 25–37, 2018.

[7] L. R. Munive-Hernández, "Predicción espacial de incendios forestales usando aprendizaje máquina," 2021.

[8] A. Baddeley *et al.*, "Analysing spatial point patterns in R," in *Workshop notes version*, vol. 3, 2008.

[9] D. J. Daley and D. Vere-Jones, *An introduction to the theory of point processes: volume II: general theory and structure*. Springer Science & Business Media, 2007.

[10] A. Baddeley, I. Bárány, and R. Schneider, "Spatial point processes and their applications," *Stochastic Geometry: Lectures Given at the CIME Summer School Held in Martina Franca, Italy, September 13–18, 2004*, pp. 1–75, 2007.

[11] A. Baddeley, E. Rubak, and R. Turner, *Spatial point patterns: methodology and applications with R*. CRC Press, 2015.

[12] E. Gabriel, F. Rodriguez-Cortes, J. Coville, J. Mateu, and J. Chadoeuf, "Mapping the intensity function of a non-stationary point process in unobserved areas," *Stochastic Environmental Research and Risk Assessment*, pp. 1–17, 2022.

[13] A. Villanueva-Morales, "Modified pseudo-likelihood estimation for Markov random fields with Winsorized Poisson conditional distributions," 2008.

[14] E. Gabriel, B. S. Rowlingson, and P. J. Diggle, "stpp: an R package for plotting, simulating and analyzing Spatio-Temporal Point Patterns," *Journal of Statistical Software*, vol. 53, pp. 1–29, 2013.

[15] N. Cressie, *Statistics for spatial data*. John Wiley & Sons, 1991.

[16] P. J. Diggle, *Statistical analysis of spatial and spatio-temporal point patterns*. CRC Press, 2013.

[17] J. Møller and R. P. Waagepetersen, *Statistical inference and simulation for spatial point processes*. Chapman and Hall/CRC, 2003.

[18] R. S. Bivand, E. J. Pebesma, V. Gómez-Rubio, and E. J. Pebesma, *Applied spatial data analysis with R*, vol. 747248717. Springer Science & Business Media, 2008.

[19] B. D. Ripley, "Modelling spatial patterns," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 2, pp. 172–192, 1977.

[20] J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan, *Statistical analysis and modelling of spatial point patterns*. John Wiley & Sons, 2008.

[21] A. Baddeley and R. Turner, "spatstat: An R package for analyzing spatial point patterns," *Journal of Statistical Software*, vol. 12, no. 6, pp. 1–42, 2005.

[22] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*. Springer Science & Business Media, 2016.

[23] A. Trapletti and K. Hornik, *tseries: Time Series Analysis and Computational Finance*, 2022. R package version 0.10-52.

[24] E. Gabriel and P. J. Diggle, "Second-order analysis of inhomogeneous spatio-temporal point process data," *Statistica Neerlandica*, vol. 63, no. 1, pp. 43–51, 2009.

[25] B. M. Taylor, T. M. Davies, B. S. Rowlingson, and P. J. Diggle, "lgcp: an r package for inference with spatial and spatio-temporal log-gaussian cox processes," *Journal of Statistical Software*, vol. 52, pp. 1–40, 2013.

[26] R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2022.