Spatial Point Processes Applied to the Study of Forest Fires in Portugal[1]

Paula Sequeira Pereira
ESTSetubal-IPS/CEAUL, paula.pereira@estsetubal.ips.pt

Kamil Feridun Turkman
DEIO-FCUL/CEAUL, kfturkman@fc.ul.pt

Abstract: The aim of this work is to analyse the behaviour of forest fires in Portugal using statistical techniques applied to spatial point processes. We present a short overview on the most commonly used summary statistics for spatial point processes under homogeneity and inhomogeneity assumptions. The data set consist of records of 6295 forest fires larger than 100 hectares, observed in Portugal during the years 1975 through 2005.

Keywords: spatial point process, K-function, marks

1 Introduction

Forest fires are a major environmental problem in Portugal. In the past few years thousands of hectares of forest have been destroyed. The aim of this work is to analyse the behaviour of forest fires in Portugal using statistical techniques applied to spatial point processes. With this analysis we intend to investigate whether the forest fires occur randomly, in clusters or in some regular pattern and we examine if the marked spatial point pattern depends on forest fires size.

This analyse is conducted in the context of a preliminary analysis of forest fires in Portugal, which is part of the general objective of modelling the location and the forest fires sizes by an adequate marked spatio-temporal point processes.

2 Materials and Methods

The point pattern under investigation consists of satellite imagery records of 6295 forest fires larger than 100 hectares, observed in Portugal during the years 1975 through 2005, acquired annually after the end of the summer fire season.

A conventional starting point for the analysis of a spatial point process is to investigate the hypothesis of complete spatial randomness (CSR). A process is CSR when we have a homogeneous Poisson point process, i.e. the intensity is constant,

---

[1] The research was supported by FCT/OE projects and SFRH/PROTEC/67394/2010 PhD grant.

and the events are independents of each other and have the same propensity to be found at any locations. If the CSR hypothesis is rejected, then must be a tendency towards clustering (events occur in closely spaced groups) or regularity (events more spaced than under CSR).

A popular tool to describe departures from CSR is the Ripley's K-function, $K(r) = \lambda^{-1}E\left[\text{number of events within distance } r \text{ of an arbitrary event}\right]$, where λ is the intensity of the point process (the number of events per unit area). Under CSR, the Ripley's K-function is simply $K(r) = \pi r^2$. Comparing the shape of ours Ripley's K-function relative to the shape of the Ripley's K-function in the case of CSR provides valuable information on the point process distribution. A Ripley's K-function that deviates from CSR can indicate that events interact or have some effect on each other, but it can also indicate that exists a trend in the pattern (the intensity of the process must not be constant across the region). By using an inhomogeneous K-function to analyse the data, is removed the assumption of an underlying homogeneous point process. The inhomogeneous K-function has the same interpretation as the homogeneous Ripley's K-function, except that the intensity of events is no longer constant but depends on the locations of the events.

In general, is common to use the L-function, which is defined as $L(r) = \sqrt{K(r)/\pi}$. Under CSR, $L(r) = r$. So we can use the line through the origin as a reference and it is simple to detect clustering or regularity by graphing $L(r) - r$ against the distance r (as in Figure 2).

The fires are characterized not only by their position but also by its size (area burned), which can be interpreted as a mark. So in addition to detection of clustering among points, the relationship between marks and between points and marks are investigated with mark correlation function defined by Stoyan (Illian et al (2008)) and **E** and **V** functions defined by Schlather et al (2004). The aim of the mark correlation function is to find out whether the marks are correlated and the aim of the **E** and **V** functions is to find out whether marks and locations are correlated (whether marks depend on local point density). If this last hypothesis is not rejected then the process is simplified greatly because the point pattern and the marks could be investigated separately.

We used the statistical software `R` and functions in the spatial point processes library `spatstat` to compute the various results.

3 Results

We start the analysis of the forest fires in a purely spatial context, so we drawn a map of the locations of the forest fires and the corresponding map of the intensity for the whole period, 1975 to 2005. Figure 1 shows that larger forest fires occur in the centre and south of Portugal but the majority of the fires are in the north of Portugal. Almost in whole country the fires are less than 0.05 fires per km^2, but in the north the highest values of the intensity are achieved, 0.2 fires per km^2.

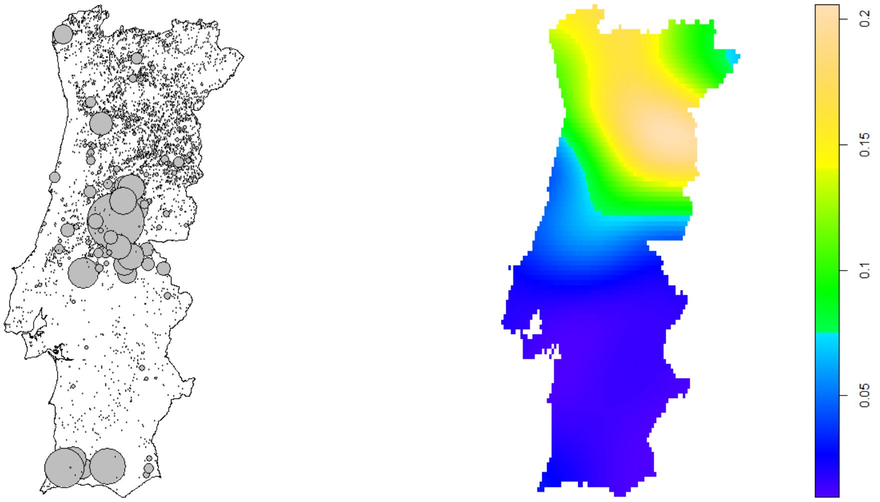

Figure 1: *Left*: Point plot of the locations of the forest fires (the circles are proportional to the area burned); *Right*: Kernel estimation of the intensity function

We compute the homogeneous and inhomogeneous L-function for all the period, 1975 to 2005. As we can see in Figure 2, there is a clear departure from CSR towards clustering. But the homogeneous L-function is overestimating the amount of clustering present in the point pattern. When we compute the inhomogeneous L-function$-r$ the distance between the estimated L-function$-r$ and the upper envelope reduces a lot. However, the inhomogeneous L-function shows evidence of clustering with a radius approximately of 50 km.

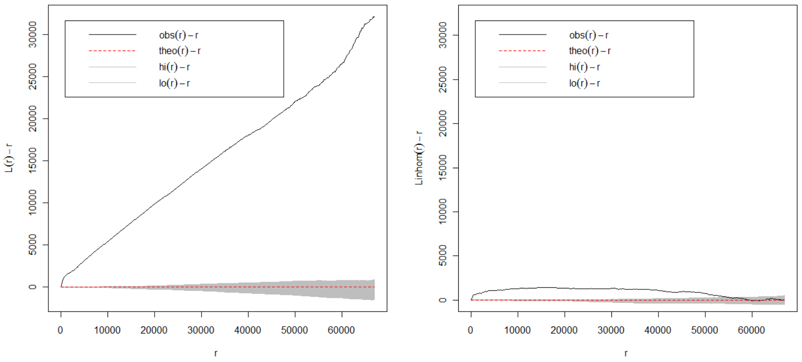

Figure 2: *Left*: Homogeneous L-function$-r$; *Right*: Inhomogeneous L-function$-r$. (In all graphics pointwise envelope under CSR and r in meters)

The mark correlation function for the area burned of the forest fires is shown in Figure 3. The shape of the empirical mark correlation function reveals that the marks do not appear to be correlated. The **E** and **V** functions indicate that the model does not appear to belong to the random field model, i.e., does not appear to be independent of the unmarked point process.

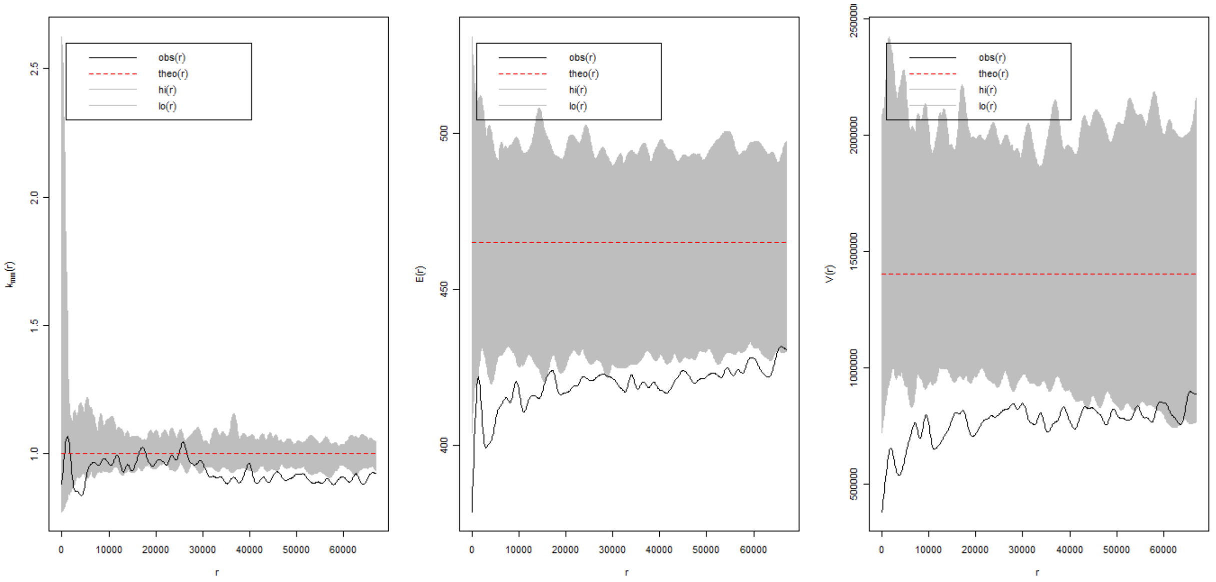

Figure 3: *Left*: Mark correlation function; *Centre*: **E** function; *Right*: **V** function. (In all graphics pointwise envelope under CSR and r in meters)

4 Concluding remarks

From the modelling point of view, the information about the behaviour of homogeneous and inhomogeneous L-function is important because it suggests that trend should be included in a model for forest fires and show that forest fires events occur in clusters, indicating spatial dependence or interaction.

The marks, area burned, do not appear to be correlated. The hypothesis of the marks to be a random field independent of the unmarked point process appears to be rejected. The point pattern and the marks should not be investigated separately.

References

Illian J., Penttinen A., Stoyan H., Stoyan D. (2008) *Statistical analysis and modelling of spatial point patterns*, Wiley. Statistics in Practice.

Schlather M., Ribeiro P., Diggle P. (2004) Detecting dependence between marks and locations of marked point processes. *J. R. Statist. Soc., Ser. B*, 66, 79-93.