Continuous testing for spatial point processes

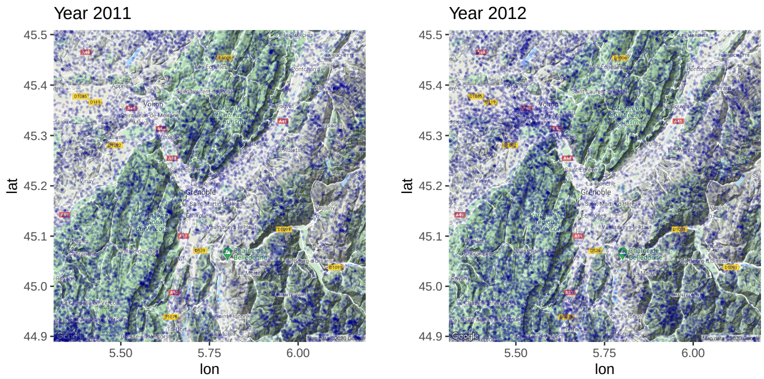
Master research internship proposal, 2020-2021

Scientific advisors	Julien Chevallier (`Julien.Chevallier1@univ-grenoble-alpes.f`) Jean-François Coeurjolly (`Jean-Francois.Coeurjolly@univ-grenoble-alpes.fr`)
Where?	Laboratory Jean Kuntzmann, Department DATA IMAG Building, 700, avenue centrale38401 Saint Martin d'Hères

Keywords	Spatial point processes; multiple testing; intensity estimation.

Prerequisites	strong background in statistics, probability and programming skills (R).

Abstract and motivation. Could we affirm that the distribution of lightning strikes around Grenoble in 2011 or 2012 is inhomogeneous in space? If yes, could we define (with statistical guarantees) regions which are responsible for this departure? Which regions cause differences in intensities when we compare 2011 and 2012? These are the typical questions, this internship intends to tackle.



Description. Spatial point processes are stochastic models which describes random events in space. Such models appear in a wide variety of applications: to model location of trees in a forest, disease locations in epidemiology, epicenters of earthquakes in seismology, etc. When we are faced with such data, the first summary statistics or question we want to address is the intensity function, a function which locally measures the probability to observe a point (a lightning strike in our case) in the vicinity of any

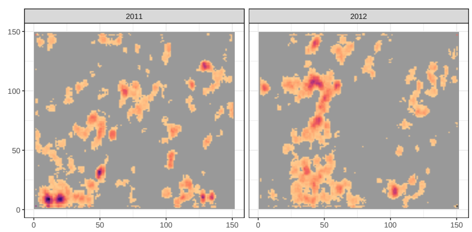


Figure 1: Map of p-values in 2011 and 2012, obtained from the scan statistics. Gray values are p-values larger than 10%.

point u in the observation domain, usually denoted by W. When this function is not constant accross space, we say the point process is inhomogeneous (see [1,3] for a review on spatial point pattern analysis and main feature characteristics for such models).

Several ways are available to test whether the intensity is homogeneous in space or not. An old one and still largely used is the scan statistics ([4]), which basically compares through a test statsitics the number of point falling into a ball centered at $u \in G \subset W$ with radius r, say, with the expected number of points if the process were homogeneous. Here, G denotes a set of grid points. Then, one obtains a map of test statistics $(T(u), u \in G)$ from which we usually obtain a map of p-values $(p(u), u \in G)$, see Figure 2 as an illustration. Of course if G is a large discrete set, we are faced with a multiple testing procedure for which the familywise error rate or false discovery rate (see [2]) have to be considered. But is this control valid if one changes G? or if G becomes more and more dense to yield the observation domain? In the latter, the problem becomes a continuous testing problem.

It turns out that [4] have proposed multiple testing procedures for such a continuous testing problem for temporal point processes. Moreover, the authors have assumed explicitly that the point pattern comes from a Poisson point process (which is the baseline process modeling random events without any intereaction between them).

Goal of this internship. (1) extend the work [4] to spatial inhomogeneous Poisson point processes; (2) implement the methodology (within the R statistical software), judge its efficiency through a simulation study and apply it to the ligthning strikes questions addressed abobe; (3) understand how far the methodology and results can be extended to more general spatial point processes.

Perspective. This research proposal may lead to a PhD Thesis on a subject more or less related to this internship.

References.

[1] Baddeley, A., Rubak, E., & Turner, R. (2015). Spatial point patterns: methodology and applications with R. CRC press.

[2] Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological), 57(1), 289-300.

[3] Coeurjolly, J. F., & Lavancier, F. (2019). Understanding spatial point patterns through intensity and conditional intensities. In Stochastic Geometry (pp. 45-85). Springer, Cham.

[4] Kulldorff, M. (1997). A spatial scan statistic. Communications in Statistics-Theory and methods, 26(6), 1481-1496.

[5] Picard, F., Reynaud-Bouret, P., & Roquain, E. (2018). Continuous testing for Poisson process intensities: a new perspective on scanning statistics. Biometrika, 105(4), 931-944.