Part 2: Point Processes

This week:
- Examples
- 3.4.1 and 3.4.2

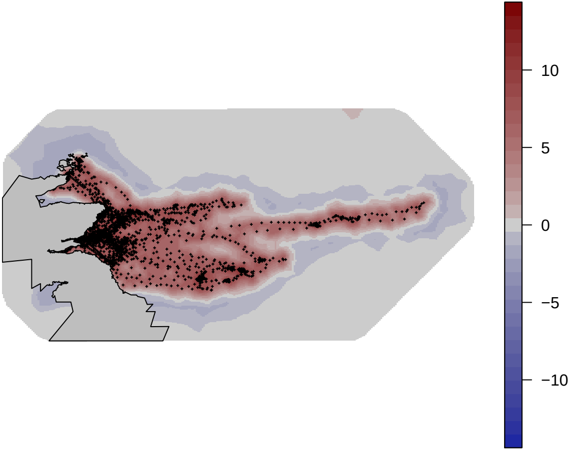
Next week:
- Simulations + more models (3.4.3 + revisit ++).

Where and when

- Tuesdays: room 734 Central Building 2 (February 14, March 7 and 14, April 6)
- Thursdays: room 743, Central Building 2 (February 9, **NOT 16**, March 9, 16, April 6)
- Friday: room 656, February 17

Grey Seals

(Presented by Haakon Bakka 1st lecture)



Avalanges crossing roads in Sogn

From Jostein Ballestad's Master Thesis:

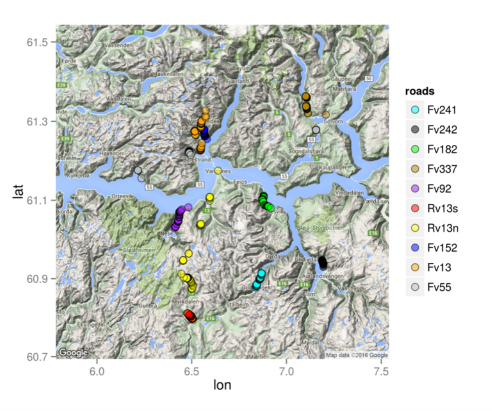
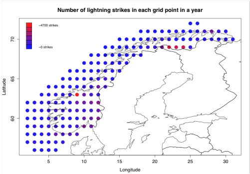


Figure 2.2: Sognefjorden with registered avalanches indicated by dots. The different colors represent different roads.

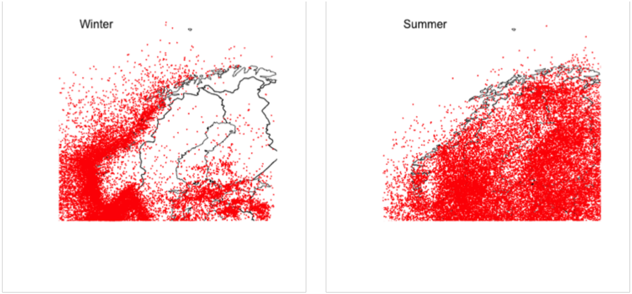
Lightning strikes over Norway From Siri Sofie Eids's project thesis.

Figure 2.3: Number of lightning strikes in each grid point in a year.



2.1 Data and explanatory analysis for strikes

Each observation in the dataset is the number of lightning strikes within an area of dimension 40×40 km^2 (N) centered in a grid point s with a certain latitude and longitude, within $\pm$ 0.5 hours of time t. There are a total of 155 different grid points, (see figure 2.1). All in all there are 602,789 observations.



(a) October-March (b) April-September

Figure 2.4: Seasonal variations in lightning strikes. Both figures show random samples of 50,000 lightning strikes from the specified months over a 10 year period.

Longleaf pine trees

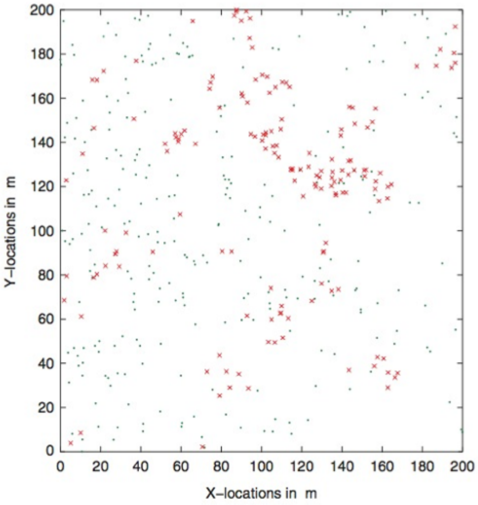


Figure 4.15 Spatial locations of longleaf pine trees in a 4-ha study area in the Wade Tract, Thomas County, Georgia, USA; the $m_A = 271$ adult longleaf pine trees are shown as green dots, and the $m_S = 159$ subadult longleaf pine trees are shown as red crosses.

Spatial Point Processe (SPP)

From Cressie and Wikle:

4.3 SPATIAL POINT PROCESSES

A spatial point process is a stochastic process governing the location of events (equivalently, points) $\{\mathbf{s}_i\}$ in some set $D_s \subset \mathbb{R}^d$, where the number of such events in D_s is also random (e.g., Diggle, 2003). In the simplest case, the

and we only consider:

Only *simple* spatial point processes in $\mathbb{R}^d$ (i.e., almost surely, either no event or a single event occurs at any point) will be considered. We characterize the

Poisson Point Process

from Walpole et al (textbook TMA4245)

Properties of Poisson Process

1. The number of outcomes occurring in one time interval or specific regions is independent of the number that occurs in any other disjunct time interval or region or space. (no memory)
2. The probability that a single outcome will occur during a very short time interval or in a small region is proportional to the size of the region and does not depend on the number of outcomes occurring outside this interval or region.
3. The probability that more then one outcome occurs in such a small time interval or fall in such a small region is negligible.

Poisson point process

An important, but simple, example is the *Poisson point process* Z, for which

$$Z(A)|\lambda^o \sim Poi(\lambda^o|A|), \qquad A \subset D_s, \qquad (4.159)$$

where $\lambda^o > 0$ is a parameter of the Poisson point process, and recall that $|A|$ is the d-dimensional volume of A. More details on this spatial point process are given in Section 4.3.1.

Sampling from Poisson point process

Trees in $1km \times 1km$ domain, with intensity $\lambda^0 = 15 trees/km^2$

-
-
-
-
-
-

ds denote a small region located at $\mathbf{s}$ with volume $|ds|$. Then the *first-order intensity* function of the point process $Z(\cdot)$ is defined as

$$\lambda(\mathbf{s}) \equiv \lim_{|ds|\to 0} E(Z(ds))/|ds|, \qquad \mathbf{s} \in D_s, \tag{4.160}$$

provided the limit exists. Hence,

$$E(Z(A)) = \int_A \lambda(\mathbf{s})\, ds, \qquad A \subset D_s. \tag{4.161}$$

Sampling from inhomogenious Poisson point process

Trees in $1km \times 1km$ domain, with intensity $\lambda(s) = f(s) trees/km^2$

-
-
-
-
-
-

Kernel estimation of intensity

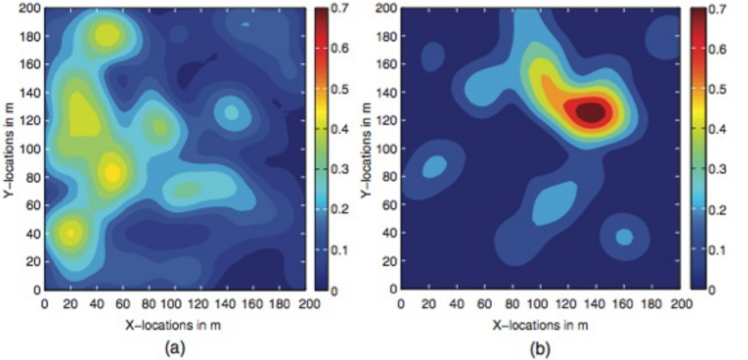
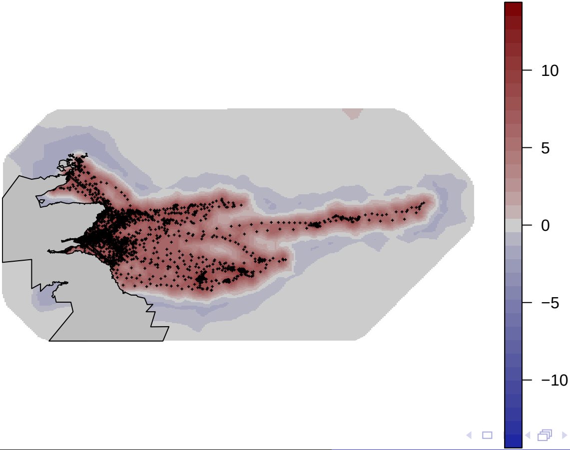


Figure 4.16 (a) Contour plot of estimated first-order intensity of $m_A = 271$ adult longleaf pine trees in the 4-ha study area in the Wade Tract; bandwidth $b = 30$ m. **(b)** Contour plot of estimated first-order intensity of $m_S = 159$ subadult longleaf pine trees in the same study area as **(a)**; bandwidth $b = 30$ m.

Estimation of intensity

We often want to estimate the $\lambda(s)$ from data (points).
Seals (from log-Gaussian Cox model):



Hierachical statistical models (HM)

Data model: $[Z|Y,\theta]$

Process model: $[Y|\theta]$

Parameter model: $[\theta]$

Gaussian Random Field (GRF) and GRF

The random field $Y(s)$, $s \in D$ (f.ex. in R^2) is a Gaussian Random field if, for any n, and any set of locations $s_1, s_2, \ldots, s_n$, all finite collections $(Y(s_1), Y(s_2), \ldots Y(s_n))$ are multivariate Normal distributed.

Need to specify the covariance function
$C_Y(s_1, s_2) = cov(Y(s_1), Y(s_2))$