Model for Random Union of Discs

Jesper Møller *) and Kateřina Helisová **)  

*) Aalborg University (Denmark)  
**) Czech Technical University/Charles University in Prague  

5th May 2008
Outline

1. Describing model
2. Simulation
3. Power tessellation of a union of discs
4. Markov property of the model
5. Data
6. Statistical inference
7. Nowadays work
**Point processes**

**Definition** Consider $N$ the system of locally finite subsets of $\mathbb{R}^d$ with the $\sigma$-algebra $\mathcal{N} = \sigma\left(\{x \in N : \#(x \cap A) = m\} : A \in \mathcal{B}, m \in \mathbb{N}_0\right)$. A point process $X$ defined on $\mathbb{R}^d$ is a measurable mapping from some probability space $(\Omega, \mathcal{F}, P)$ to $(N, \mathcal{N})$.

**Definition** A locally finite, diffusion measure $\mu$ on $\mathcal{B}$ satisfying $\mu(A) = \mathbb{E}X(A)$ for all $A \in \mathcal{B}$ is called the intensity measure.

**Definition** If there exists a function $\rho(x)$ for $x \in \mathbb{R}^d$ such that $\mu(A) = \int_A \rho(x) dx$, then $\rho(x)$ is called the intensity function.

**Definition** A point process is called finite point process if $\mu(\mathbb{R}^d) < \infty$. 
Poisson point process

Definition The Poisson process $Y$ is the process which satisfies:

- for any finite collection $\{A_n\}$ of disjoint sets in $\mathbb{R}^d$, the numbers of points in these sets, $Y(A_n)$, are independent random variables,
- for each $A \subset \mathbb{R}^d$ such that $\mu(A) < \infty$, $Y(A)$ has Poisson distribution with parameter $\mu(A)$, i.e. $P[Y(A) = k] = \frac{\mu(A)^k}{k!}e^{-\mu}$, where $\mu$ is the intensity measure.
Point process given by the density with respect to Poisson process

Let $Y$ be the Poisson process with an intensity measure $\mu$.

For $F \in \mathcal{N}$, denote $\Pi(F) = P(Y \in F)$.

**Definition** A point process $X$ is given by density $f$ with respect to the Poisson process $Y$ if

$$P(X \in F) = \int_F f(x) \Pi(dx).$$
Model

Denoting $b = b(z, r)$ a disc with centre in $z \in \mathbb{R}^2$ and radius $r \in (0, \infty)$, we have a process of discs $\bigcup b_i = \bigcup b(z_i, r_i)$. Then, we identify $b$ with the point $x = (z, r)$ in $\mathbb{R}^2 \times (0, \infty)$ and the process of discs $\bigcup b_i = \bigcup b(z_i, r_i)$ with a point process on $\mathbb{R}^2 \times (0, \infty)$.

The reference point process: A Poisson process $Y$ (so that the reference Boolean model is the random set given by the union of discs in $Y$) with intensity measure $\rho(z) \, dz \, Q(dr)$ on $S \times (0, \infty)$, where $S \subset \mathbb{R}^2$ is bounded.

Model: The process $X$ absolutely continuous with respect to the reference Poisson process $Y$, and given by density $f(x)$ for a finite configurations $x = \{x_1, \ldots, x_n\}$.
Exponential family model

General form of the density:

\[ f_\theta(x) = \exp(\theta \cdot T(U_x)) / c_\theta \]

Set \( T = (A, L, \chi, N_{ic}, N_h, N_{bv}) \), where

- \( A = A(U_x) \) ...the area
- \( L = L(U_x) \) ...the perimeter
- \( \chi = \chi(U_x) \) ...the Euler-Poincaré characteristic
- \( N_h = N_h(U_x) \) ...the number of holes
- \( N_{ic} = N_{ic}(U_x) \) ...the number of isolated cells
- \( N_{bv} = N_{bv}(U_x) \) ...the number of boundary vertices,

i.e. the density is of the form

\[
f_\theta(x) = \frac{1}{c_\theta} \exp \left( \theta_1 A(U_x) + \theta_2 L(U_x) + \theta_3 \chi(U_x) + \theta_4 N_h(U_x) + \theta_5 N_{ic}(U_x) + \theta_6 N_{bv}(U_x) \right).
\]
Some simulated results

A power tessellation of a realization of the reference Poisson process with \( Q \) the uniform distribution on the interval \([0, 2]\), \( \rho(u) = 0.2 \) on a rectangular region \( S = [0, 30] \times [0, 30] \), and \( \rho(u) = 0 \) outside \( S \) (left) and \( A \)-interaction model with parameters \( \theta_1 = 0.1 \) (middle), resp. \( \theta_1 = -0.1 \) (right).
Some simulated results

$(A, L, N_{cc})$-interaction process, where $N_{cc}(U_x)$ is the number of connected components, with parameters $(0.6, -1, 1)$ (left), $(0.6, -1, 2)$ (middle) and $(0.6, -1, 5)$ (right).
Papangelou conditional intensity

**Definition** For finite $x \subset S \times (0, \infty)$ and $v \in S \times (0, \infty) \setminus x$, *Papangelou conditional intensity* is defined as

$$\lambda_\theta(x, v) = \frac{f_\theta(x \cup \{v\})}{f_\theta(x)}.$$ 

Denoting $W = A, L, \ldots$ and defining $W(x, v) = W(x \cup v) - W(x)$, we get

$$\lambda_\theta(x, v) = \exp(\theta_1 \cdot A(x, v) + \theta_2 \cdot L(x, v) + \cdots + \theta_6 \cdot N_{bv}(x, v)).$$
MCMC algorithm

1. Suppose that in time $t$, we have a configuration $x_t = \{x_1, \ldots, x_n\}$

2. Proposal in time $t + 1$:
   
   (a) with probability $1/2$, the proposal is $x_t \cup \{x_{n+1}\}$
   
   i. we accept the proposal with probability $\min\{1; h(x_t, x_{n+1})\}$

   and set $x_{t+1} = x_t \cup \{x_{n+1}\}$

   ii. else we set $x_{t+1} = x_t$

   (b) else, the proposal is $x_t \setminus \{x_i\}$

   i. we accept the proposal with probability $\min\{1; 1/h(x_t \setminus \{x_i\}, x_i)\}$

   and set $x_{t+1} = x_t \setminus \{x_i\}$

   ii. else $x_{t+1} = x_t$

where $h(x_t, x_{n+1}) = \lambda_\theta(x_t, x_{n+1}) \frac{|S|}{\rho \cdot (n+1)}$

and $h(x_t \setminus \{x_i\}, x_i) = \lambda_\theta(x_t \setminus \{x_i\}, x_i) \frac{|S|}{\rho \cdot n}$. 
Power tessellation of a union of discs

Assume a union of discs $\mathcal{U} = \bigcup_{i} b_i$ in the general position.

For each disc $b_i$ ($i \in I$) with ghost sphere $s_i$, let $s_i^+ = \{(y_1, y_2, y_3) \in s_i : y_3 \geq 0\}$ denote the corresponding upper hypersphere.

For $u \in b_i$, let $y_i(u)$ denote the unique point on $s_i^+$ those orthogonal projection on $\mathbb{R}^2$ is $u$.

Define

$$C_i = \{y_i(u) : u \in b_i, \|u - y_i(u)\| \geq \|u - y_j(u)\| \text{ for } u \in b_j, j \in I\}.$$  

Denote $B_i$ the orthogonal projection of $C_i$ on $\mathbb{R}^2$.

**Definition** The system $\mathcal{B}$ of all sets $B_i$ is called a *power tessellation of a union of discs*. 
Power tessellation of a union of discs

Left: A configuration of discs in general position. Middle: The upper hemispheres as seen from above. Right: The power tessellation of the union of discs.
Usefulness of power tessellation in MCMC algorithm

• $\chi(U_x) = N_c - N_{ie} + N_{iv}$

• Calculation of $A(U_x)$: instead of

$$A(U_x) = \sum_i A(b_i) - \sum_{\{i_1, i_2\}} A(b_{i_1} \cap b_{i_2}) \ldots (-1)^{n+1} \sum_{\{i_1, \ldots, i_n\}} A(b_{i_1} \cap \ldots \cap b_{i_n})$$

we use

$$A(U_x) = \sum_i B_i$$

• Analogously calculation of $L(U_x)$
Markov property

Recall that Papangelou conditional intensity is defined as 
\[ \lambda_\theta(x, v) = \frac{f_\theta(x \cup \{v\})}{f_\theta(x)}. \]

**Definition** The \( T \)-interaction process is said to be Markov with respect to a reflexive relation \( \sim \) (called *neighbourhood*) if \( \lambda_\theta(x, v) \) depends on \( x \) only through \( \{u \in x : u \sim v\} \), i.e. the neighbours in \( x \) to \( v \).

**Remark:** If the relation \( \sim \) depends on the configuration \( x \), the process is called nearest neighbour Markov process.

**Hammersley-Clifford theorem** A function \( f \) is the density of a Markov process if and only if there exists a function \( \phi \) such that 
\[ f(x) = \prod_{y \subset x} \phi(y) \] and \( \phi(y) = 1 \) whenever there exist \( y_i, y_j \in y \) such that \( y_i \not\sim y_j \).
Markov property - in terms of the connected components

The density is of the form

\[ f_\theta(x) = \frac{1}{C_\theta} \prod_{K \in \mathcal{K}(U_x)} \exp(\theta_1 A(K) + \theta_2 L(K) + \theta_3 \chi(K) + \theta_4 N_h(K) + \theta_5 N_{ic}(K) + \theta_6 N_{bv}(K)) \]

where \( \mathcal{K}(U_x) \) is the set of connected components of \( U_x \).

Hammersley-Clifford theorem

\[ \downarrow \]

the \( T \)-interaction process is a connected component Markov point process.
Data

Heather dataset first presented in (Diggle, 1981). The image shows the presence of heather (calluna vulgaris) (indicated by black) in a $10 \times 20$ m region at Jädraås, Sweeden.
Diggle’s model

Diggle modeled the presence of heather by a stationary random-disc Boolean model with the centres given by a stationary Poisson process and the radii i.i.d., independent of the centres, with Weibull distribution.

**Problem:** He observed that his fitted model generates more separate patches than in the data.

↓

Our model
Statistical inference - problems

Denote $\mathbf{Y}$ the random set and $W$ the observation window.

The statistical inference is complicated by the fact that

(i) we do not observe the individual discs but only their union within $W$,

(ii) in practice the grains may only approximately be discs and usually only a digital image is observed and the resolution makes it difficult to identify circular structures,

(iii) there occurs a problem with edge effects when $\mathbf{Y}$ expands outside $W$. 

Modified model

Recall

\[ f_\theta(x) = \frac{\exp(\theta \cdot T(U_x))}{c_\theta}. \]

Original statistic

\[ T = (A, L, \chi, N_h, N_{ic}, N_{bv}) \]

is replaced by

\[ T = (A, L, N_{cc}, N_h), \]

i.e. the density is considered in the form

\[ f_\theta(x) = \frac{1}{c_\theta} \exp(\theta_1 A(U_x) + \theta_2 L(U_x) + \theta_3 N_{cc}(U_x) + \theta_4 N_h(U_x)). \]
Solving of the edge effects problem

Split $X$ into $X^{(a)}$, $X^{(b)}$, $X^{(c)}$ corresponding to discs belonging to connected components of $\mathcal{U}_X$ which are respectively (a) contained in the window $W$, (b) intersecting both $W$ and its complement $W^c$, (c) contained in $W^c$.

Let $x^{(b)}$ denote a realization of $X^{(b)}$, i.e. $x^{(b)}$ is a finite configuration of discs such that $K$ intersects both $W$ and $W^c$ for all $K \in \mathcal{K}(\mathcal{U}_{x^{(b)}})$.

It holds that conditional on $X^{(b)} = x^{(b)}$, we have that $X^{(a)}$ and $X^{(c)}$ are independent, and the conditional distribution of $X^{(a)}$ depends only on $x^{(b)}$ through $V = W \cap \mathcal{U}_{x^{(b)}}$. 
Solving edge effects problem

Illustrating possible realizations of $X^{(a)}$ (the full circles), $X^{(b)}$ (the dashed circles), and $X^{(c)}$ (the dotted circles).
Solving edge effects problem

Heather dataset without the components intersected by the boundary of the observation window.

**Remark:** For the data, $A(V)/A(W) = 0.2734$ when $A(Y \cap W)/A(W) = 0.5014 \Rightarrow$ loss of information.
Reference process

A realization of the reference Boolean model with the intensity $\rho = 2.45$ and $R$ following the restriction of $N(0.26, 0.16^2)$ to the interval $[0, 0.50]$. 
Estimates of parameters

Denote $f_\theta(x) = h_\theta(x)/c_\theta$ (i.e. $h_\theta(x) = \exp(\theta \cdot T(U_x))$ is the unnormalized density).

For an observation $x$, the log likelihood function is given by

$$l(\theta) = \log h_\theta(x) - \log c_\theta$$

and the score function

$$u(\theta) = d l(\theta)/d\theta = d \log h_\theta(x)/d\theta - d \log c_\theta/d\theta.$$

In the exponential family case, we have

$$l(\theta) = \theta \cdot T(U_x) - \log c_\theta \text{ and } u(\theta) = T(U_x) - \mathbb{E}_\theta T(U_x).$$

Problem: $c_\theta$ and $\mathbb{E}_\theta T(U_x)$ have no explicit expression.
Estimates of parameters

For fixed $\theta_0$, the log likelihood ratio

$$l(\theta) - l(\theta_0) = \log(h_\theta(x)/h_{\theta_0}(x)) - \log(c_\theta/c_{\theta_0})$$

can be approximated by

$$l(\theta) - l(\theta_0) = \log(h_\theta(x)/h_{\theta_0}(x)) - \log \frac{1}{n} \sum_{m=0}^{n-1} h_\theta(Y_m)/h_{\theta_0}(Y_m),$$

where $Y_m$ are realizations from $f_{\theta_0}(x)$ obtained from MCMC simulations.

Recall

$$u(\theta) = \frac{dl(\theta)}{d\theta}$$

and denote

$$j(\theta) = -\frac{du(\theta)}{d\theta}.$$
Estimates of parameters

Newton-Raphson method for maximising the log likelihood:

1. Set $\hat{\theta}^{(0)} = \theta_0$;

2. $(k+1)$-th iteration is given by

$$\hat{\theta}^{(k+1)} = \hat{\theta}^{(k)} + u_{\theta_0,n}(\hat{\theta}^{(k)}) \cdot j_{\theta_0,n}(\hat{\theta}^{(k)})^{-1},$$

where

$$u_{\theta_0,n}(\hat{\theta}^{(k)}) = T(\mathcal{U}_x) - \mathbb{E}_{\hat{\theta}^{(k)}, \theta_0,n} T(\mathcal{U}_X)$$

$$= T(\mathcal{U}_x) - \frac{\sum_{m=0}^{n-1} T(\mathcal{U}_{Y_m}) h_{\hat{\theta}(m)}(Y_m)/h_{\theta_0}(Y_m)}{\sum_{m=0}^{n-1} h_{\hat{\theta}(k)}(Y_m)/h_{\theta_0}(Y_m)}$$

and

$$j_{\theta_0,n}(\hat{\theta}^{(k)}) = \text{Var}_{\hat{\theta}^{(k)}, \theta_0,n} T(\mathcal{U}_X).$$
Estimates of parameters from the data

Using the heather data without the components intersected by the boundary of the observation window, we have

\[ T = (A, L, N_{cc}, N_h) = (45.6(m^2), 204(m), 32, 2). \]

For the reference Boolean model we obtain

\[ T = (A, L, N_{cc}, N_h) = (42.7(m^2), 220.8(m), 111, 6). \]
Estimates of parameters from the data

Simulation:

1. Simulations from $\theta_0 = (0, 0, 0, 0)$

2. After iterations 1 000 000 and 100 steps od N-R method we have $\hat{\theta} = (-1.70, 0.38, -0.73, -0.67)$

3. Simulations from $\theta_0 = (-1.70, 0.38, -0.73, -0.67)$:

n. Repeating the same procedure seven times, we get $\hat{\theta} = (-4.35, 1.02, -2.23, 0.89)$

Characteristics of the model with parameters $(-4.35, 1.02, -2.23, 0.89)$:

$$T = (A, L, N_{cc}, N_h) = (46.82(m^2), 199.99(m), 32, 2).$$
Simulated process

A realization of the $T$-interaction model with the parameters $(-4.35, 1.02, -2.23, 0.89)$. 
Nowadays work

- Test the model against Boolean model and against some other sub-models
- Check the model
- Distribution of radii
Thank you for your attention!