Bayesian analysis of spatial point processes in the neighbourhood of Voronoi networks

Øivind Skare, Jesper Møller and Eva B. Vedel Jensen

Bayesian analysis of spatial point processes in the neighbourhood of Voronoi networks

Bayesian analysis of spatial point processes in the neighbourhood of Voronoi networks

Øivind Skare,[*] Jesper Møller,[†] Eva B. Vedel Jensen[‡]

ABSTRACT: A model for an inhomogeneous Poisson process with high intensity near the edges of a Voronoi tessellation in 2D or 3D is proposed. The model is analysed in a Bayesian setting with priors on nuclei of the Voronoi tessellation and other model parameters. An MCMC algorithm is constructed to sample from the posterior, which contains information about the unobserved Voronoi tessellation and the model parameters. A major element of the MCMC algorithm is the reconstruction of the Voronoi tessellation after a proposed local change of the tessellation. A simulation study and examples of applications from biology (animal territories) and material science (alumina grain structure) are presented.

Keywords: Bayesian inference, Delaunay tessellation, inhomogeneous point processes, Markov chain Monte Carlo, Poisson process, Voronoi tessellation.

1 Introduction

Spatial point processes in the neighbourhood of a reference structure are often observed. A diverse set of examples at very different scales are copper deposits in the neighbourhood of lineaments (Berman 1986), gold coins near Roman roads (Hodder & Orton 1976), specific tree species along rivers in the rain forest (Valencia et al. 2004), galaxies at the boundary of cosmic voids (Icke & Van de Weygaert 1987, Van de Weygaert & Icke 1989, Van de Weygaert 1994), pores at the boundary of grains (Karlsson & Liljeborg 1994) and animal latrines near territory boundaries (Blackwell 2001, Blackwell & Møller 2003). In most cases the reference structure is unknown. The object of the statistical analysis of point patterns of this type may either be to describe the distribution of the point pattern or to reconstruct the reference structure or perhaps both.

In the present paper, we propose a model for spatial point processes in the neighbourhood of the edges of a Voronoi tessellation. The model is analysed in a Bayesian setting. The idea of using Voronoi tessellations for constructing spatial processes in a Bayesian setting is not new. In the seminal paper Green (1995), reversible jump Markov chain Monte Carlo has been developed and applied to image segmentation (subdivision of a digital image into homogeneous regions) via Voronoi tessellations. Heikkinen & Arjas (1998) and Heikkinen & Arjas (1999) studied a

[*]Department of Biostatistics, University of Oslo, Norway
[†]Department of Mathematical Sciences, Aalborg University, Denmark
[‡]Department of Mathematical Sciences, University of Aarhus, Denmark

nonparametric Bayesian modelling framework for inhomogeneous Poisson processes where the intensity function is piecewise constant on Voronoi cells. In Møller & Skare (2001), Voronoi tessellations have been used in a Bayesian setting for reservoir modelling.

The model discussed in the present paper can be used for modelling inhomogeneous point processes. Recently, models for inhomogeneous point processes have been suggested by several authors, cf. Baddeley et al. (2000), Jensen & Nielsen (2001), Hahn et al. (2003) and references therein. Although some of these models may be adapted to include explanatory variables such as the distance to the reference structure the models are not particularly focused on point patterns lying near a reference structure. We propose a model for such a point pattern, obtained by perturbing points on the Voronoi edges with intensity ρ, using a standard normal distribution with independent coordinates and common variance σ^2 for all coordinates. Bayesian inference is developed in both 2D and 3D. A major step in the posterior simulation from the model is the local change of the Voronoi tessellation at each iteration. It is easy to modify the simulation algorithm in case another point intensity function is of interest, including the one suggested in Blackwell (2001).

The present paper is organized as follows. In Section 2, the suggested model is described, including likelihood and prior assumptions, while posterior simulation from the model is discussed in Section 3. A simulation study is presented in Section 4. Bayesian analysis of two data sets from biology and material science, respectively, is described in Section 5. Section 6 concludes with a discussion.

2   The model

In this section we specify the various steps of a Bayesian hierarchical model construction for a Poisson process with an intensity function which is concentrated in the neighbourhood of the edges of a Voronoi tessellation. We consider first the case where the Voronoi tessellation is known.

2.1   Voronoi tessellations

Before specifying our model we need to introduce some terminology and notation and to recall a few properties of Voronoi tessellations. For further background material on Voronoi tessellations, see Møller (1994) and Okabe et al. (2000).

Let $\bar{B} \subset \mathbf{R}^d$ be a bounded closed convex d-dimensional set where $d \geq 2$, and let $\mathbf{y} = \{y_i\} \subset \bar{B}$ be a finite set of points called nuclei. In our application examples, $d = 2$ or 3 and $\bar{B}$ is a rectangle or a box containing an observation window B. The Voronoi cell $C(y_i|\mathbf{y})$ with nucleus $y_i \in \mathbf{y}$ and restricted to $\bar{B}$ consists of all points in $\bar{B}$ that is not closer to another nucleus $y_j \in \mathbf{y} \backslash y_i$:

$$C(y_i|\mathbf{y}) = \left\{ s \in \bar{B} : \|y_i - s\| \leq \|y_j - s\| \text{ for all } y_j \in \mathbf{y} \right\}, \qquad y_i \in \mathbf{y},$$

where $\|\cdot\|$ denotes Euclidean distance. The Voronoi tessellation $\mathcal{V}(\mathbf{y})$ generated by $\mathbf{y}$ and restricted to $\bar{B}$ is the collection of Voronoi cells $C(y_i|\mathbf{y})$, $y_i \in \mathbf{y}$. For an example of planar Voronoi tessellations, see Figure 1 below.

Henceforth, to avoid "degenerated cases", see Møller (1994), we assume that the nuclei satisfy the following conditions:

- $n(\mathbf{y}) \geq d$ where $n(\mathbf{y})$ denotes the number of nuclei;

- $\mathbf{y}$ is in general quadratic position, that is, no $k + 1$ nuclei lie on a $k - 1$ dimensional affine subspace of $\mathbf{R}^d$, $k = 2, \dots, d$, and no $d + 2$ nuclei lie on the boundary of a sphere.

These conditions and the assumptions imposed on $\bar{B}$ ensure the following. Any non-void intersection of $j \geq 2$ Voronoi cells is a bounded closed convex $(d + 1 - j)$-dimensional set (this set is empty if $j \geq d+2$). In particular any non-void intersection of d Voronoi cells is a finite closed line segment of positive length called a Voronoi edge. We denote the union of edges by $E(\mathbf{y})$ and the total length of $E(\mathbf{y})$ by $L(\mathbf{y})$. Note that $\emptyset \neq E(\mathbf{y}) \subset \bar{B}$ and $0 < L(\mathbf{y}) < \infty$. An endpoint of an edge is called an interior or a boundary vertex of $\mathcal{V}(\mathbf{y})$ depending on whether it belongs to the interior or boundary of $\bar{B}$. The interior vertices are the non-void intersections of any $d + 1$ Voronoi cells, and so exactly $d + 1$ edges emerge at an interior vertex. In fact, under the prior model for $\mathbf{y}$ specified in Section 2.4, with probability one, $\mathbf{y}$ is finite and in general quadratic position, and only one edge emerges at any boundary vertex. Note that in the spatial case $d = 3$, a non-void intersection between two Voronoi cells is a bounded closed convex two-dimensional set called a face.

For the simulation algorithm in Section 3 we need the Delaunay tessellation, the dual tessellation to $\mathcal{V}(\mathbf{y})$. A Delaunay cell is a d-dimensional simplex (i.e. a triangle if $d = 2$ or a tetrahedron if $d = 3$) given by the convex hull of $d + 1$ nuclei so that the closed ball containing these nuclei in its boundary does not contain any further nuclei. Thus the union of Delaunay cells is the convex hull of $\mathbf{y}$ which in general will be strictly contained in $\bar{B}$. We consider later the extended Delaunay tessellation $\mathcal{T}(\mathbf{y})$ given by adding $d + 1$ fixed dummy points $\mathbf{v} = \{v_0, \dots, v_d\} \subset \mathbf{R}^d \setminus \bar{B}$: Each Delaunay cell in $\mathcal{T}(\mathbf{y})$ corresponds to $d + 1$ points (or "nuclei") from $\mathbf{w} = \mathbf{v} \cup \mathbf{y}$ such that the closed ball containing these $d + 1$ points in its boundary contains no other points from $\mathbf{w}$. Here it is required that

- $\mathbf{w}$ is in general quadratic position;

- the original Delaunay cells defined by $\mathbf{y}$ are contained in $\mathcal{T}(\mathbf{y})$;

- $\bar{B}$ is contained in the convex hull of $\mathbf{w}$;

- the dummy points do not influence the Voronoi tessellation in $\bar{B}$, i.e. $\mathcal{V}(\mathbf{y} \cup \mathbf{v}) = \mathcal{V}(\mathbf{y})$.

The latter requirement is obtained by choosing each v_i such that the shortest distance from v_i to any point in $\bar{B}$ is larger than $\sup\{\|\varepsilon - \eta\| : \varepsilon, \eta \in \bar{B}\}$. For example for $\bar{B} = [0, \delta] \times [0, \delta]$, we chose $v_0 = -(\delta, \delta)$, $v_1 = (2\delta, -\delta)$ and $v_2 = (0.5\delta, 2.5\delta)$. As explained later in Section 3, the important fact is that from $\mathcal{T}(\mathbf{y})$ we can easily obtain $\mathcal{T}(\mathbf{y} \cup \{\xi\})$ and hence $\mathcal{V}(\mathbf{y} \cup \{\xi\})$ when adding a point $\xi \in \bar{B}$ such that $\mathbf{y} \cup \{\xi\}$ is in general quadratic position.

2.2 Data distribution

Suppose that our data is $\mathbf{x} \cap B = \mathbf{x}_B$ where $\mathbf{x}$ is a Poisson process defined on $\mathbf{R}^d$ and observed within a window $B \subset \mathbf{R}^d$ with d-dimensional Lebesgue measure $|B| \in (0, \infty)$. Below we construct the intensity function of this Poisson process by a two-step procedure such that the intensity function near a given set $E \subset \mathbf{R}^d$ is high.

First, we introduce a latent Poisson process $\mathbf{z}$ with intensity measure Λ concentrated on E and such that $0 < \Lambda(E) < \infty$ (ensuring that with probability one $\mathbf{z}$ is finite and with a positive probability $\mathbf{z}$ is non-empty). Second, the point process $\mathbf{x} = \{x_i\}$ is obtained from the point process $\mathbf{z} = \{z_i\}$ by i.i.d. perturbations $\{\varepsilon_i\}$ which are independent of $\mathbf{z}$:

$$x_i = z_i + \varepsilon_i.$$

Assume that ε_i has density h with respect to Lebesgue measure on $\mathbf{R}^d$. Then $\mathbf{x}$ is a Poisson process with intensity function

$$\chi_E(x|\Lambda, h) = \int_E h(x - z)\, \Lambda(\mathrm{d}z),$$

see the "Displacement Theorem" in Kingman (1993) or Section 3.3.1 in Møller & Waagepetersen (2003). Note that the total intensity of $\mathbf{x}$ is given by $\int_{\mathbf{R}^d} \chi_E(x|\Lambda, h)\, \mathrm{d}x = \Lambda(E)$.

Specifically, we assume that $E = E(\mathbf{y})$ is specified as in Section 2.1 where $\bar{B} \supseteq B$ is "sufficiently large" such that boundary effects can be ignored, see the discussion of Figure 1 below. Furthermore, we assume that $\mathbf{z}$ is homogeneous on E with intensity $\rho > 0$ (i.e. $\Lambda(\mathrm{d}z) = \rho\, \mathrm{d}z$, where $\mathrm{d}z$ denotes Lebesgue measure on E), and h is the density of a radially symmetric d-dimensional normal distribution $\mathcal{N}_d(0, \sigma^2 I_d)$ with variance $\sigma^2 > 0$. Then $\Lambda(E) = \rho L(\mathbf{y}) \in (0, \infty)$ with probability one, as required.

We denote the intensity function $\chi_E(x|\Lambda, h)$ by

$$\chi_E(x|\rho, \sigma^2) = \frac{\rho}{(2\pi\sigma^2)^{d/2}} \int_E \exp\left(-\frac{\|x - z\|^2}{2\sigma^2}\right)\, \mathrm{d}z \tag{1}$$

which can be simplified as follows. For any edge $e = [u, v] \in E$, let p_e denote the orthogonal projection onto the line containing e. By Pythagoras, $\|x - z\|^2 = \|x - p_e(x)\|^2 + \|p_e(x) - z\|^2$, whereby (1) reduces to

$$\chi_E(x|\rho, \sigma^2) = (2\pi\sigma^2)^{-(d-1)/2} \rho \sum_{e \in E} i(\sigma^2, e, x) \tag{2}$$

with

$$
\begin{aligned}
i(\sigma^2, e, x) = \exp\left(-\frac{1}{2\sigma^2}\|x - p_e(x)\|^2\right) &\left\{ \left[\Phi\left(\frac{d_u}{\sigma}\right) + \Phi\left(\frac{d_v}{\sigma}\right) - 1\right] \mathbf{1}[p_e(x) \in e] \right. \\
&\left. + \left[\Phi\left(\frac{\max(d_u, d_v)}{\sigma}\right) - \Phi\left(\frac{\min(d_u, d_v)}{\sigma}\right)\right] \mathbf{1}[p_e(x) \notin e] \right\}
\end{aligned}
$$

where

$$d_u = \|u - p_e(x)\|, \qquad d_v = \|v - p_e(x)\|.$$

Figure 1 shows an example of the intensity function $\chi_E(x|\rho, \sigma^2)$ and the Voronoi tessellation $\mathcal{V}(\mathbf{y})$ restricted to a square $B = [0, 10]^2$, when $(\sigma, \lambda) = (0.4, 0.16)$ and the size of $\bar{B}$ is varied. Boundary effects are clearly seen for the two smallest choices of $\bar{B}$, while for the two largest choices the intensity functions and Voronoi tessellations are essentially equal. Note that the value of $\chi_E(x|\rho, \sigma^2)$, when x is close to an interior Voronoi vertex, is determined by the following: a) all edges sharing the vertex will contribute to the intensity; b) on the other hand, $i(\sigma^2, e, x)$ is decreasing when $\|p_e(x) - m\|$ or $\|x - p_e(x)\|$ are increasing, where $m = (u + v)/2$ is the midpoint of e. In Figure 1, a) is more important than b) as $\chi_E(x|\rho, \sigma^2)$ takes its highest values for x near the vertices of $\mathcal{V}(\mathbf{y})$. In higher dimensions $d \geq 3$, a) is even more pronounced, since more edges emerges from interior vertices.
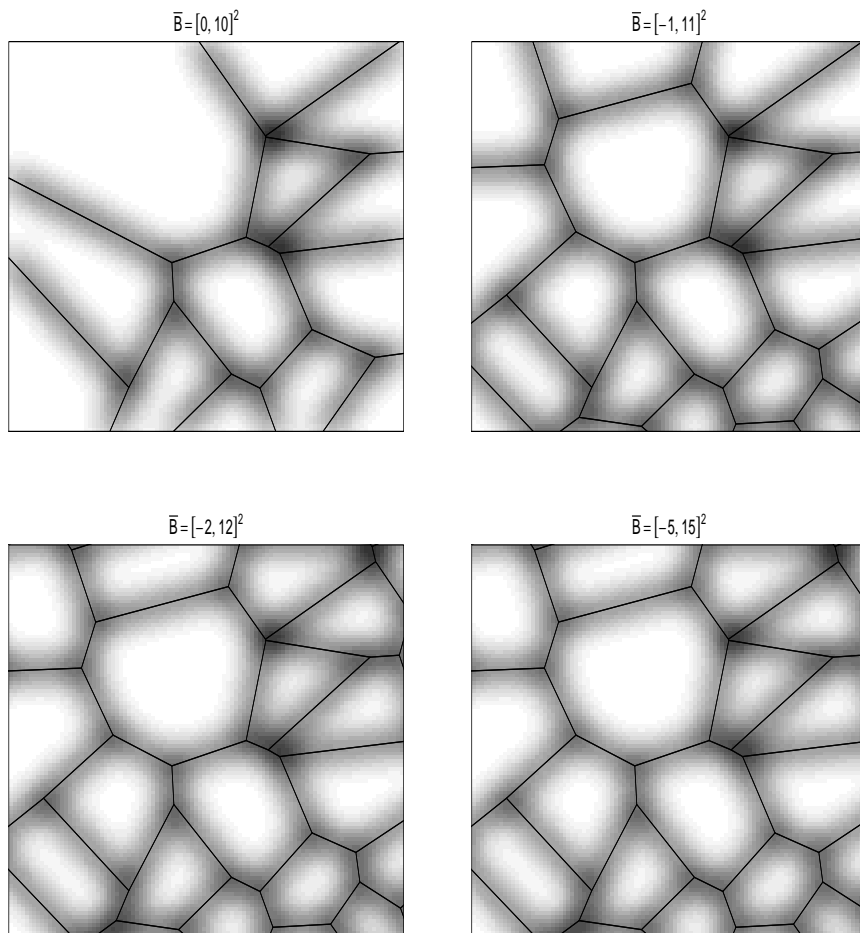

Figure 1: Grey scale plot of the intensity function $\chi_E(x|\rho, \sigma^2)$ for given Voronoi tessellations restricted to $B = [0, 10]^2$, where $\sigma = 0.4$ and $\bar{B}$ is varying. The nuclei generating the Voronoi tessellation is a simulation of a homogeneous Poisson process on $\bar{B}$ with intensity $\lambda = 0.16$. The nuclei for the smaller sets $\bar{B}$ are the restriction to these sets of the nuclei generated in the largest set $\bar{B} = [-5, 15]^2$.

2.3 Likelihood

The likelihood based on the data $\mathbf{x}_B$ turns out to be intractable because it involves the calculation of the integral $\int_B \chi_E(x|\rho, \sigma^2) \, dx$. For the Bayesian analysis we use therefore the "full" likelihood based on $\mathbf{x}$ and treat $\mathbf{x}_{B^c} = \mathbf{x} \setminus \mathbf{x}_B$ as "missing data". It turns out that there is no need to include $\mathbf{z}$ as "missing data".

Specifically, let μ denote the distribution of a finite Poisson process on $\mathbf{R}^d$ with strictly positive intensity function k so that $\int_{\mathbf{R}^d} k(x) \, dx < \infty$. Then the "full" likelihood is given by

$$l(\rho, \sigma^2, \mathbf{y}|\mathbf{x}) = \left[\prod_{x_i \in \mathbf{x}} \frac{\chi_E(x_i|\rho, \sigma^2)}{k(x_i)} \right] \exp(-\rho L(\mathbf{y})), \tag{3}$$

since this times the constant $\exp\left(\int k(x) \, dx \right)$ is equal to the density of $\mathbf{x}$ with respect to μ, cf. Proposition 3.8 in Møller & Waagepetersen (2003). The choice of k plays no importance in the sequel.

2.4 Prior assumptions

In cases where we have no immediately available prior knowledge about the parameters ρ and σ^2, and possibly also no prior knowledge about $\mathbf{y}$, we seek a relatively non-informative prior distribution. Thus, we assume that $\rho, \sigma^2, \mathbf{y}$ are independent, with $\rho \sim U(0, \rho_{\max})$ and $\sigma \sim U(0, \sigma_{\max})$, where $\rho_{\max} > 0$ and $\sigma_{\max} > 0$ are known, and $U(a, b)$ denotes the uniform distribution on the interval (a, b). If $\mathbf{y}$ is unknown, $\mathbf{y}$ is assumed to be a homogeneous Poisson process on $\bar{B}$ with intensity λ, where we have conditioned on that $n(\mathbf{y}) \geq d$ (see Section 2.1), and where $\lambda \sim U(0, \lambda_{\max})$, with $\lambda_{\max} > 0$ known. Then (with probability one) $\mathbf{y}$ is finite and in general quadratic position, cf. Møller (1994). Note that

$$\pi(\rho) \propto \mathbf{1}[0 < \rho < \rho_{\max}], \;\; \pi(\sigma^2) \propto \frac{1}{\sigma}\mathbf{1}[0 < \sigma < \sigma_{\max}], \;\; \pi(\lambda) \propto \mathbf{1}[0 < \lambda < \lambda_{\max}], \tag{4}$$

where $\mathbf{1}[\cdot]$ denotes the indicator function. We generally choose large values for $\lambda_{\max}$, $\rho_{\max}$ and $\sigma_{\max}$ such that these limits do not influence the posterior distributions. For later use, if $\bar{\mu}$ denotes the distribution of a unit rate Poisson process on $\bar{B}$,

$$\pi(\mathbf{y}|\lambda) = \frac{\lambda^{n(\mathbf{y})} \exp((1 - \lambda)|\bar{B}|)}{1 - \exp(-\lambda|\bar{B}|) \sum_{i=0}^{d-1} (\lambda|\bar{B}|)^i/i!} \tag{5}$$

is the conditional density of a Poisson process with respect to $\bar{\mu}$, when we condition on both the intensity λ and the restriction $n(\mathbf{y}) \geq d$.

2.5 Posterior

By Bayes theorem, the posterior density is

$$\pi(\rho, \sigma^2, \mathbf{x}_{B^c}|\mathbf{x}_B, \mathbf{y}) \propto \pi(\rho)\pi(\sigma^2)l(\rho, \sigma^2, \mathbf{y}|\mathbf{x}) \tag{6}$$

when **y** is known, and

$$\pi(\rho, \sigma^2, \lambda, \mathbf{y}, \mathbf{x}_{B^c} | \mathbf{x}_B) \propto \pi(\rho)\pi(\sigma^2)\pi(\mathbf{y}|\lambda)\pi(\lambda)l(\rho, \sigma^2, \mathbf{y}|\mathbf{x}) \qquad (7)$$

when **y** is unknown, where the terms on the right hand side are given by (3)–(5). In both cases of (6) and (7) we need to use Markov chain Monte Carlo (MCMC) simulations as explained in detail below.

3 Posterior simulation

We presume that the reader has some familiarity to MCMC algorithms, see e.g. Gilks et al. (1996). For simulation from (6) or (7) we use a fixed scan Metropolis within Gibbs algorithm (also called a hybrid algorithm), where in each scan ρ, σ^2 and $\mathbf{x}_{B^c}$, and also **y** and λ when these are unknown (the case of (7)), are updated in turn. Briefly, the updates for ρ, σ^2 and λ are respectively simple Gibbs, Metropolis random walk, and independence sampling updates; the updates of $\mathbf{x}_{B^c}$ are easily generated by a thinning procedure; and the updates of **y** are of birth, death or move types. Details are given below.

3.1 Updates for ρ, σ^2 and λ

The full conditional for ρ is the restriction of a Gamma distribution:

$$\rho \,|\, (\sigma^2, \lambda, \mathbf{y}, \mathbf{x}) \ \sim \Gamma(n(\mathbf{x}) + 1, L(\mathbf{y})), \ \text{with the constraint that } \rho < \rho_{\max}.$$

Here we use a Gibbs update, using rejection sampling (i.e. sampling from the Gamma distribution until the constraint is satisfied).

The full conditional

$$\pi(\sigma^2 | \rho, \lambda, \mathbf{y}, \mathbf{x}) \propto \pi(\sigma^2)l(\rho, \sigma^2, \mathbf{y}|\mathbf{x})$$

is not a standard distribution, so instead of Gibbs sampling we use Gaussian random walk updates for σ^2: If $(\rho, \sigma^2, \lambda, \mathbf{y})$ is the current state, we generate a normal proposal $\sigma_*^2 \sim \mathcal{N}(\sigma^2, \kappa)$ and accept σ_*^2 with probability

$$\min\left\{1, \frac{\pi(\sigma_*^2)}{\pi(\sigma^2)} \prod_{x_i \in \mathbf{x}} \frac{\chi_{E(\mathbf{y})}(x_i|\rho, \sigma_*^2)}{\chi_{E(\mathbf{y})}(x_i|\rho, \sigma^2)}\right\}$$

and retain σ^2 otherwise. To get some idea of how to choose κ, we ran Markov chains for different values of κ and compared their estimated first-order autocorrelations. This suggested that the optimal value of κ corresponds to an acceptance probability that is slightly above 0.40, in close agreement with Roberts & Rosenthal (2001) and Gelman et al. (1996).

The full conditional

$$\pi(\lambda|\rho, \sigma^2, \mathbf{y}, \mathbf{x}) \propto \pi(\mathbf{y}|\lambda)\pi(\lambda)$$

7

is a $\Gamma(n(\mathbf{y}) + 1, |\bar{B}|)$ distribution, if we ignore the constraint that $n(\mathbf{y}) \geq d$. Therefore, we use independence sampling updates for λ: If $(\rho, \sigma^2, \lambda, \mathbf{y})$ is the current state, we generate a proposal $\lambda_* \sim \Gamma(n(\mathbf{y}) + 1, |\bar{B}|)$ and accept λ_* with probability

$$\min\left\{1, \frac{1 - \exp(-\lambda|\bar{B}|)\sum_{i=0}^{d-1}(\lambda|\bar{B}|)^i/i!}{1 - \exp(-\lambda_*|\bar{B}|)\sum_{i=0}^{d-1}(\lambda_*|\bar{B}|)^i/i!}\mathbf{1}[\lambda_* < \lambda_{\max}]\right\}$$

and retain λ otherwise. In our application examples, the acceptance probability is close to 1.

3.2 Updates for $\mathbf{x}_{B^c}$

We simulate from the conditional distribution of $\mathbf{x}_{B^c}$ given $(\rho, \sigma^2, \lambda, \mathbf{y}, \mathbf{x}_B)$ by the following thinning procedure which is based on the fact that conditional on $(\rho, \sigma^2, \lambda, \mathbf{y})$, the Poisson processes $\mathbf{x}_B$ and $\mathbf{x}_{B^c}$ are independent.

- Simulate a homogeneous Poisson process $\mathbf{z}'$ on $E(\mathbf{y})$ with intensity ρ.

- For each $z \in \mathbf{z}'$, generate a random variable $x(z)$ from $\mathcal{N}_d(z, \sigma^2 I_d)$, where all the $x(z)$ are conditionally independent given $\mathbf{z}'$.

- Return the collection of those points $x(z)$ with $x(z) \in B^c$.

3.3 Updates for $\mathbf{y}$

Since $\mathbf{y}$ is a point process with a varying number of points, its full conditional

$$\pi(\mathbf{y}|\rho, \sigma^2, \lambda, \mathbf{x}) \propto \pi(\mathbf{y}|\lambda)l(\rho, \sigma^2, \mathbf{y}|\mathbf{x})$$

is more complicated than the full conditionals above. As in Geyer & Møller (1994), we let each update of $\mathbf{y}$ consists of either a birth, death or move proposal, and the proposal is accepted with probability $\min\{1, H\}$, where H is the Hastings ratio.

Specifically, suppose that $(\rho, \sigma^2, \mathbf{y}, \lambda)$ is the current state and we want to update $\mathbf{y}$, where $n(\mathbf{y}) \geq d$ as required. A birth or a death or a move proposal occurs with probability q, q or $1 - 2q$, respectively, where $0 < q \leq 0.5$ is a user-specified parameter. In case of a birth proposal $\mathbf{y} \to \mathbf{y} \cup \{\xi\}$, say, the new point ξ is uniformly distributed on $\bar{B}$ and $H = R(\mathbf{y}; \xi)$, where

$$R(\mathbf{y}; \xi) = \frac{\lambda|\bar{B}|}{n(\mathbf{y}) + 1}\left[\prod_{x_i \in \mathbf{x}}\frac{\chi_{E(\mathbf{y} \cup \{\xi\})}(x_i|\rho, \sigma^2)}{\chi_{E(\mathbf{y})}(x_i|\rho, \sigma^2)}\right]\exp\left(\rho(L(\mathbf{y}) - L(\mathbf{y} \cup \{\xi\}))\right). \quad (8)$$

In case of a death proposal $\mathbf{y} \to \mathbf{y} \setminus \{\eta\}$, say, the point η is chosen uniformly at random from $\mathbf{y}$ and $H = 1/R(\mathbf{y} \setminus \{\eta\}; \eta)$ (setting $H = 0$ if $n(\mathbf{y}) \leq d$). In case of a move proposal $\mathbf{y} \to (\mathbf{y} \setminus \{\eta\}) \cup \{\xi\}$, say, we use a Metropolis random walk update: choose η uniformly at random from $\mathbf{y}$ and propose to replace this by ξ, drawn from $\mathcal{N}_d(\eta, \sigma^2_{\text{move}} I_d)$ (see Sections 4 and 5 for the choice of σ_{move}). Then $H = R(\mathbf{y} \setminus \{\xi\}; \eta)/R(\mathbf{y} \setminus \{\eta\}; \xi)$ (setting $H = 0$ if $\xi \notin \bar{B}$).

3.4 Construction of the Voronoi tessellation

The critical update is that for $\mathbf{y}$ (and partly also that for σ^2), since a birth or death proposal can cause a locally large change of $E(\mathbf{y})$ and hence a low acceptance probability. We use an incremental method for building the Delaunay tessellation $\mathcal{T}(\mathbf{y})$ and thereby the Voronoi tessellation $\mathcal{V}(\mathbf{y})$, cf. Section 2.1. Briefly, for a birth proposal $\mathbf{y} \to \mathbf{y} \cup \{\xi\}$, we make first an initial tessellation of simplices by searching for the Delaunay cell $T \in \mathcal{T}(\mathbf{y})$ that contains ξ. If T has vertices $w_0, \ldots, w_d$, then we replace T by the $d + 1$ new simplices with vertices $w_0, \ldots, w_{i-1}, \xi, w_{i+1}, \ldots, w_d$, $i = 0, \ldots, d$. Second, a number of edge (if $d = 2$) or edge and face (if $d = 3$) flip operations are performed until the Delaunay property is restored, see Section 2.1 and the algorithm described in Mücke (1993). Our extension of Mücke's algorithm simply consists in including reverse operations: A death proposal $\mathbf{y} \to \mathbf{y} \setminus \{\eta\}$ also involves edge and face flip operations, going in the reverse direction of the corresponding birth proposal $\mathbf{y} \setminus \{\eta\} \to \mathbf{y}$. A move proposal $\mathbf{y} \to (\mathbf{y} \setminus \{\eta\}) \cup \{\xi\}$ involves the edge and face flip operations for first the death $\mathbf{y} \to \mathbf{y} \setminus \{\eta\}$ and next the birth $\mathbf{y} \setminus \{\eta\} \to (\mathbf{y} \setminus \{\eta\}) \cup \{\xi\}$.

4 Simulation study

We considered two simulated data sets: In each case the data $\mathbf{x}_B$ was generated using the same underlying tessellation as in Figure 1 with $\bar{B} = [-2, 12]^2$, where the nuclei of the Voronoi tessellation was a realisation of a Poisson point process with $\lambda = 0.16$. The first data set was simulated with $\sigma = 0.4$ and $\rho = 2$, the same values that produced the intensity $\chi_E(\cdot|\rho, \sigma^2)$ shown in Figure 1. The second data set was simulated with $\sigma = 0.1$, while $\rho = 2$ as before. Thus the second data set was more informative with respect to $\mathbf{y}$ than the first data set.

We used identical and independent uniform priors: λ, ρ and σ follow $U(0, 1000)$. For the updates of $\mathbf{y}$, we used the MCMC algorithm described in Section 3.3 with $q = 0.35$. Furthermore, $\sigma_{\text{move}} = 0.8$ and $\sigma_{\text{move}} = 0.14$ for the two cases, and the acceptance probabilities for a move proposal were 0.22 and 0.25, respectively. Finally, $\kappa = 0.06^2$ and $\kappa = 0.005^2$ in the random walk proposal for σ^2 for the two cases, and the acceptance probability were 0.36 and 0.27, respectively.

The posterior edge intensity surface was calculated by dividing B into 200×200 squares of equal size, and estimating by MCMC methods for each square the probability that the square contains a Voronoi edge. For the first data set the positions of the Voronoi edges showed a large degree of variability as judged from the posterior edge intensity, while in the other case the data more or less fixed the Voronoi tessellation in the posterior simulations. As expected the variability was largest at the boundary of B. Figure 2 shows the posterior edge intensity surface together with the simulated data $\mathbf{x}_B$ obtained in the first case.

A total of 5,000,000 iterations were run with a burn-in of 500,000 iterations. Trace plots of $n(\mathbf{y})$, $n(\mathbf{x})$, λ, ρ, σ^2 and $\log(\pi(\rho, \sigma^2, \lambda, \mathbf{y}, \mathbf{x}_{B^c}|\mathbf{x}_B))$ (similar to Figure 4 below, but omitted here) indicated that the chains were well mixing. The posterior distributions had heavy tails, notably for σ^2, but also for λ (and hence for $n(\mathbf{y})$) in the first case with true $\sigma^2 = 0.16$. We estimated the integrated auto-correlation
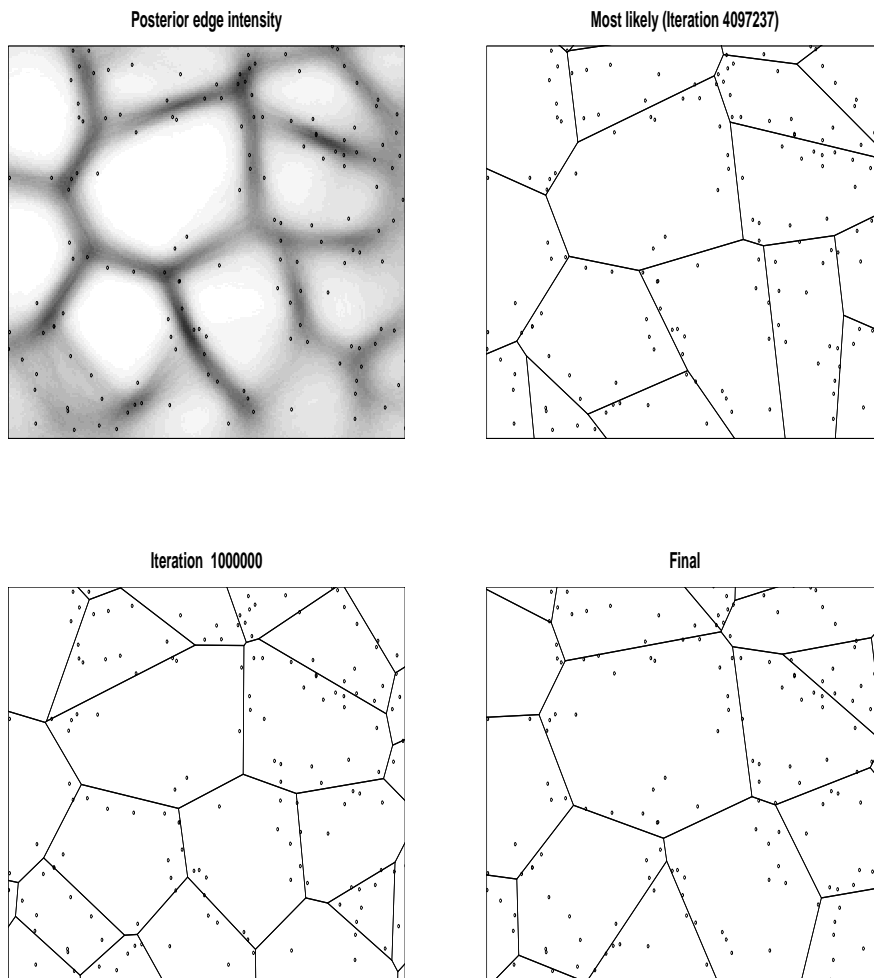
Figure 2: Simulation study. First data set, with $B = [0, 10]^2$, $\bar{B} = [-2, 12]^2$, $\sigma = 0.4$, $\rho = 2$ and $\lambda = 0.16$. Upper left figure: posterior intensity surface together with the simulated data. The other figures show realisations from the posterior: the tessellation with highest posterior probability, the tessellation at iteration 1,000,000 and the final tessellation after 5,000,000 iterations.

time (IACT) from the samples of $n(\mathbf{y})$, using the initial sequence estimator for the asymptotic variance (Geyer 1992). The estimates were IACT $= 31,400$ for the first case and IACT $= 85,500$ for the second case. Hence, as expected, the algorithm mixed fastest in the first case.

5 Examples

5.1 Reconstructing territories of badgers

Animal territories have often been modelled by Voronoi tessellations, see Covich (1976), Tanemura & Hasegawa (1980), Blackwell (2001), Blackwell & Møller (2003)

and references therein. The example in this section concerns the reconstruction of territories of badgers from the positions of latrines. We assume that these territories form a Voronoi tessellation and that the badgers create latrines near the borders of their territories. The data set consists of the locations of 124 latrines observed in Wythan Woods, Oxford, U.K, see the point pattern in the upper left panel of Figure 3. The data have previously been analysed in Blackwell (2001) and Blackwell & Møller (2003).

Posterior edge intensity **Most likely (Iteration 96112675)**

Iteration 150003375 **Final**

Figure 3: Badgers data. Upper left figure: posterior intensity surface together with the locations of 124 latrines. The other figures show realisations from the posterior: the most likely tessellation, the tessellation at iteration 150,003,375 and the final tessellation after 201,000,000 iterations.

As in Section 4, we used identical and independent priors, $U(0, 1000)$, for λ, ρ and σ. We assumed an unknown number of generating points in the Voronoi tessellation, i.e. an unknown number of territories. Figure 3 shows the posterior edge intensity and realisations from the posterior. A total of 201,000,000 iterations were run with a burn-in of 1,000,000 iterations. Figure 4 shows trace plots of $n(\mathbf{y})$, $n(\mathbf{x})$, λ, ρ, σ^2 and $\log(\pi(\rho, \sigma^2, \lambda, \mathbf{y}, \mathbf{x}_{B^c}|\mathbf{x}_B))$ as a function of iteration number.

In Figure 5, a plot of $\hat{L}(r) - r$ as a function of the distance r is shown for the observed data together with the average value of $\hat{L}(r) - r$ obtained from simulations from the posterior predictive distribution, and upper and lower 2.5% percentiles. Recall that $\hat{L}(r) = \sqrt{\hat{K}(r)/\pi}$ where $\hat{K}(r)$ is the estimated value of the K-function at r, see e.g. Stoyan et al. (1995). The samples from the posterior predictive distri-
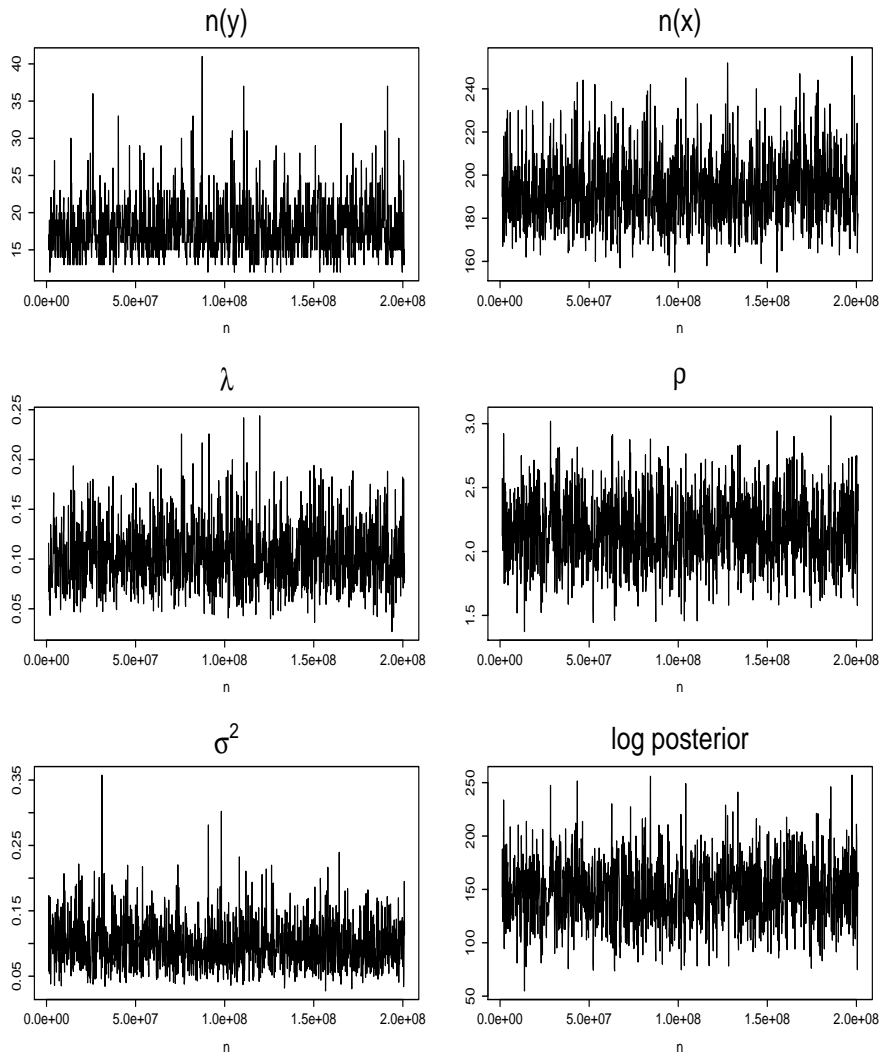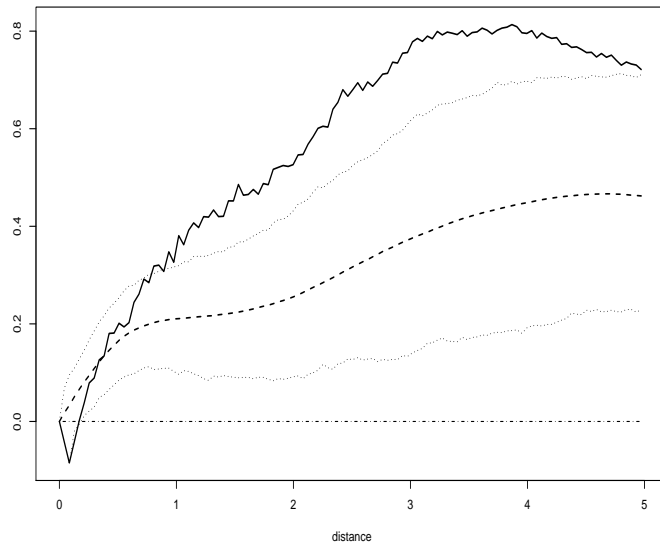
Figure 4: Badgers data. Trace plots of $n(\mathbf{y})$, $n(\mathbf{x})$, λ, ρ, σ^2 and $\log(\pi(\rho, \sigma^2, \lambda, \mathbf{y}, \mathbf{x}_{B^c}|\mathbf{x}_B))$ as a function of iteration number.

bution were taken with spacing given by the estimated IACT which was 279,290. In Figure 6, a few simulated point patterns from the posterior predictive distribution are compared with the observed point pattern.

Judged from the posterior edge intensity in Figure 3 and the simulated point patterns in Figure 6, the model captures important features of the data, but alignment of points appears to be more pronounced in the observed point pattern. This is confirmed by Figure 5, which shows that the model somewhat underestimates the degree of clustering.

Figure 5: Badgers data. Plot of $\hat{L}(r) - r$ as a function of the distance r is shown for the observed data (solid thick line), together with the average value of $\hat{L}(r) - r$ obtained from simulations from the posterior predictive distribution (dashed thick line) and corresponding upper and lower 2.5% percentiles (dotted lines). In the Poisson case, $L(r) - r = 0$, here shown as a dotted-dashed line for comparison. The distance 5 is half the width of the window in Figure 3.
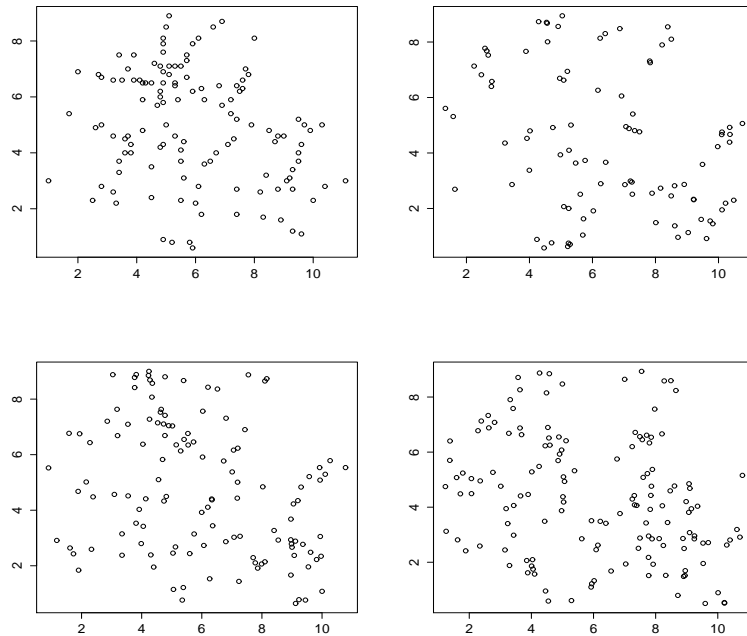

Figure 6: Badgers data. Observed point pattern (upper left corner) and simulated point patterns from the posterior predictive distribution.

5.2 Distribution of pores in translucent alumina

This example concerns the 3D distribution of pores in translucent alumina. These pores are located at the boundaries of the alumina grains. In Karlsson & Liljeborg (1994), the spatial distribution of such pores has been studied, using confocal scanning laser microscopy for the registration of 3D coordinates of centroids of pores. It was found, using summary statistics, that the pores form a clustered point pattern. Here, we investigate whether a model of the type discussed in the present paper is appropriate for the description of the spatial distribution of pores.

We consider one of the data sets analysed in Karlsson & Liljeborg (1994), consisting of the 3D coordinates of 122 pore centroids x_i in a block $B = [0, 94] \times [0, 94] \times [0, 54]$ (measures are in micro meter (μm)), see Figure 7. In the analysis, we used identical and independent priors, $U(0, 1000)$, for λ, ρ and σ. We used a standard deviation of 0.25 in the random walk proposal for σ^2. This gave an acceptance probability around 50 % for σ^2. For the updates of $\mathbf{y}$, we chose $\sigma_{\text{move}} = 2.0$. A simulation of 205,000,000 iterations was run with a burn-in of 5,000,000. Plots of $n(\mathbf{y})$, $n(\mathbf{x})$, λ, ρ, σ^2 and $\log(\pi(\rho, \sigma^2, \lambda, \mathbf{y}, \mathbf{x}_{B^c} | \mathbf{x}_B))$ as a function of iteration number were well-behaved as in the case of badgers data.
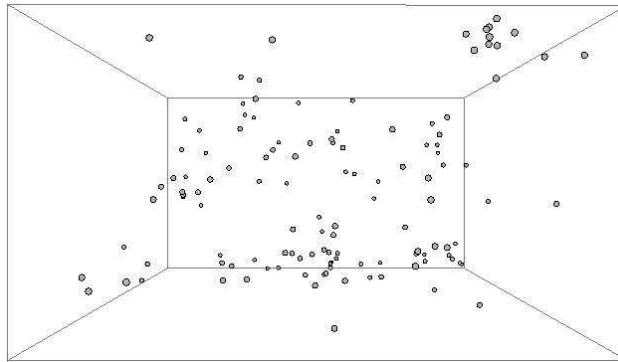

Figure 7: The alumina pore data. The 3D coordinates of the centroids of 122 pores in a 94×94×54 μm^3 block B of translucent alumina have been recorded by confocal scanning laser microscopy.

The estimated value of the posterior mean of λ is 7.14×10^{-5} μm^{-3}. Instead of specifying the intensity of cells, we may use the mean width w of a Voronoi cell. By Stoyan et al. (1995), page 330, w can be estimated by

$$w = \left[\frac{16\pi^5}{243}\right]^{1/3} \frac{\Gamma(1/3)}{5} \lambda^{-1/3} = 35.1 \mu m.$$

This estimate appears plausible from visual inspection of images of serial sections through the matrix of alumina grains, cf. Karlsson & Liljeborg (1994).

In Figure 8, a plot of $\hat{L}(r) - r$ as a function of the distance r is shown for the observed data together with the average value of $\hat{L}(r) - r$ obtained from simulations from the posterior predictive distribution, and upper and lower 2.5% percentiles. The samples from the posterior predictive distribution were taken with spacing

given by the estimated IACT which was 1,903,600. Since the points are now in 3D, we have

$$\hat{L}(r) = \left[\hat{K}(r)/(4\pi/3)\right]^{1/3}.$$

As judged from Figure 8, the observed point patterns appear to be somewhat more clustered at small distances that predicted. Simulated point patterns from the posterior distribution were also compared with the observed point pattern, using projection. No marked difference was found.

In Figure 9, a 3D illustration of the posterior edge intensity surface is shown together with the observed pore centroids.
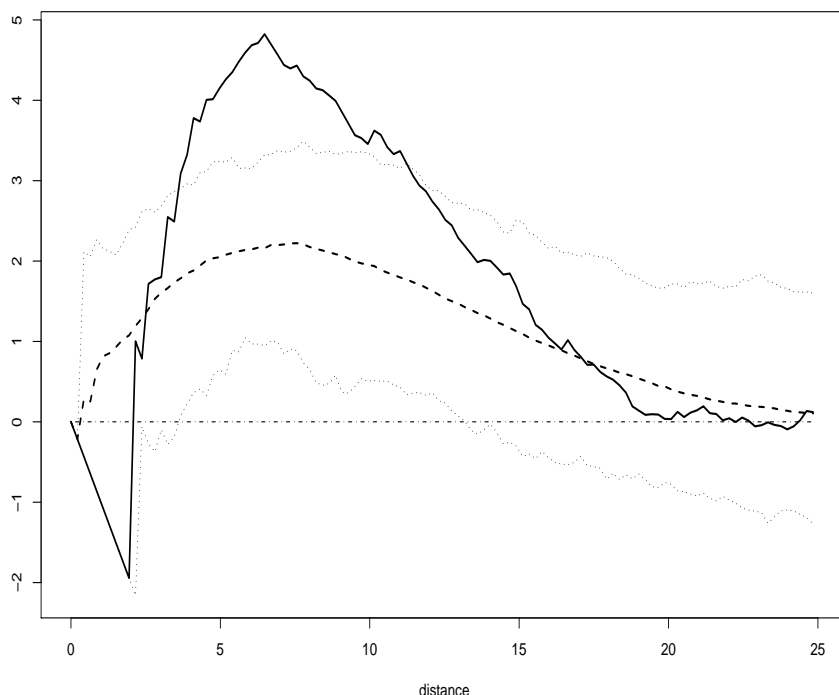

Figure 8: Pore data. Plot of $\hat{L}(r) - r$ as a function of the distance r is shown for the observed data (solid thick line), together with the average value of simulated $\hat{L}(r) - r$ obtained from simulations from the posterior predictive distribution (dashed thick line) and corresponding upper and lower 2.5% percentiles (dotted lines). In the Poisson case, $L(r) - r = 0$, here shown as a dotted-dashed line for comparison. The distance is in μm.

6   Discussion

In the analysis of both simulated and real data we have used uniform priors. If prior knowledge is available about some of the parameters in the model, non-uniform priors should be used instead for these parameters. In Blackwell (2001) and Blackwell
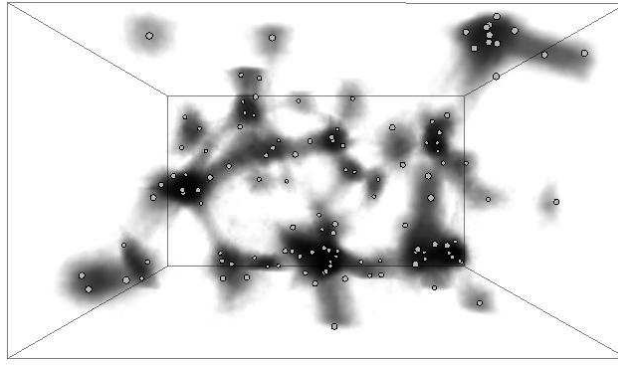
Figure 9: Pore data. The posterior intensity surface together with the observed pore data.

& Møller (2003), information about the locations of the setts (burrows) of the badgers has been used in the analysis. A total of 16 setts was identified for the data analysed in Section 5.1. The setts were used as nuclei of the Voronoi tessellation. However, as shown in Blackwell (2001) and Blackwell & Møller (2003), nuclei located away from the setts give a better fit to the data. Apart from that, the biological reasons for using the setts as **y** are somewhat unclear. For these reasons we have chosen not to include the setts in our analysis of the badger's data.

It turns out that the model considered in the present paper is of considerable interest in cosmology, cf. Van de Weygaert (2005, personal communication). It is part of our future research plans to use the developed model and its natural extensions in the study of the spatial distribution of galaxies.

7   Acknowledgements

References

Baddeley, A., Møller, J. & Waagepetersen, R. (2000), 'Non- and semi-parametric estimation of interaction in inhomogeneous point patterns', *Statist. Neerlandica* **54**, 329–350.

Berman, M. (1986), 'Testing for spatial association between a point process and another stochastics process', *Appl. Statist.* **35**, 54–62.

Blackwell, P. G. (2001), 'Bayesian inference for a random tessellation process', *Biometrics* **57**, 502–507.

Blackwell, P. G. & Møller, J. (2003), 'Bayesian analysis of deformed tessellation models', *Adv. Appl. Probab.* **35**, 4–26.

Covich, A. P. (1976), 'Analyzing shapes of foraging areas: Some ecological and economic theories', *Ann. Rev. Ecol. Syst.* **7**, 235–257.

Gelman, A., Roberts, G. O. & Gilks, W. R. (1996), Efficient Metropolis jumping rules, *in* J. M. Bernado, J. O. Berger, A. P. Dawid & A. F. M. Smith, eds, 'Bayesian Statistics 5', Oxford University Press, Oxford, pp. 599–607.

Geyer, C. J. (1992), 'Practical Markov chain Monte Carlo', **7**, 473–483.

Geyer, C. J. & Møller, J. (1994), 'Simulation procedures and likelihood inference for spatial point processes', *Scand. J. Statist.* **21**, 359–373.

Gilks, W. R., Richardson, S. & Spiegelhalter, D. J. (1996), *Markov Chain Monte Carlo in Practice*, Chapman & Hall, London.

Green, P. J. (1995), 'Reversible jump MCMC computation and Bayesian model determination', *Biometrika* **82**, 711–732.

Hahn, U., Jensen, E. B. V., Lieshout, M. N. M. & Nielsen, L. S. (2003), 'Inhomogeneous spatial point processes by location-dependent scaling', *Adv. Appl. Probab.* **35**, 319–336.

Heikkinen, J. & Arjas, E. (1998), 'Non-parametric Bayesian estimation of a spatial Poisson intensity', *Scand. J. Statist.* **25**, 435–450.

Heikkinen, J. & Arjas, E. (1999), 'Modelling a Poisson forest in variable elevations: a nonparametric Bayesian approach', *Biometrics* **55**, 738–745.

Hodder, I. & Orton, C. (1976), *Spatial Analysis in Archaeology*, Cambridge University Press, chapter 7, pp. 226–229.

Icke, V. & Van de Weygaert, R. (1987), 'Fragmenting the Universe', *Astron. Astrophys.* **184**, 16–32.

Jensen, E. B. V. & Nielsen, L. S. (2001), A review on inhomogeneous spatial point processes, *in* I. V. Basawa, C. C. Heyde & R. L. Taylor, eds, 'Selected Proceedings of the Symposium on Inference for Stochastic Processes', Vol. 37, IMS Lecture Notes - Monographs Series, Ohio, pp. 297–318.

Karlsson, L. M. & Liljeborg, A. (1994), 'Second-order stereology for pores in translucent alumina studied by confocal scanning laser microscopy', *J. Microsc.* **175**, 186–194.

Kingman, J. F. C. (1993), *Poisson Processes*, Clarendon Press, Oxford.

Møller, J. (1994), *Lectures on Random Voronoi Tessellations*, Lecture Notes in Statistics 87, Springer-Verlag, New York.

Møller, J. & Skare, Ø. (2001), 'Bayesian image analysis with coloured Voronoi tessellations and a view to applications in reservoir modelling', *Statistical Modelling* **1**, 213–232.

Møller, J. & Waagepetersen, R. P. (2003), *Statistical Inference and Simulation for Spatial Point Processes*, Chapman and Hall/CRC, Boca Raton.

Mücke, E. (1993), Shapes and Implementations in Three-Dimensional Geometry, PhD thesis, University of Illinois.

Okabe, A., Boots, B., Sugihara, K. & Chiu, S. N. (2000), *Spatial Tessellations. Concepts and Applications of Voronoi Diagrams*, second edn, Wiley, Chichester.

Roberts, G. O. & Rosenthal, J. S. (2001), 'Optimal scaling for various Metropolis-Hastings algorithms', *Statist. Sci.* **16**, 351–367.

Stoyan, D., Kendall, W. S. & Mecke, J. (1995), *Stochastic Geometry and Its Applications*, second edn, Wiley, Chichester.

Tanemura, M. & Hasegawa, M. (1980), 'Geometrical models of territory 1. Models for synchronous and asynchronous settlement of territories', *J. Theor. Biol.* **82**, 477–496.

Valencia, R., Foster, R. B., Villa, G., Condit, R., Svenning, J.-C., Hernández, C., Romoleroux, K., Losos, E., Magård, E. & Balslev, H. (2004), 'Tree species distributions and local habitat variation in the Amazon, Ecuador: large forest plot in Eastern Ecuador', *J. Ecol.* **92**, 214–229.

Van de Weygaert, R. (1994), 'Fragmenting the Universe III. The construction and statistics of 3-d Voronoi tessellations', *Astron. Astrophys.* **283**, 361–406.

Van de Weygaert, R. & Icke, V. (1989), 'Fragmenting the Universe II. Voronoi vertices as Abell clusters', *Astron. Astrophys.* **213**, 1–9.