Bayesian Estimation and Segmentation of Spatial Point Processes using Voronoï Tilings.

Simon D. Byers and Adrian E. Raftery [1]

July 29, 1997
Technical Report Number 326
Department of Statistics
University of Washington.

Abstract

We consider the problem of detecting features in spatial point processes in the presence of substantial clutter, and the problem of intensity estimation of spatial point processes. One example is the location of seismic faults from earthquake catalogs; another is the detection of minefields using reconnaissance aircraft images that identify many objects that are not mines.

Our solution uses the Voronoï tiling of a set of synthetic generating points to provide a partition of the region of interest. This partition then provides the basis for modeling the spatial point process as having piecewise constant intensity. The model that this generates is estimated using Markov chain Monte Carlo. The number of generating points can be either specified by the user or estimated; in the latter case, reversible jump MCMC is used.

In the minefield problem this method yields good detection rates and few false positives. The method provides a large amount of probabilistic information about the estimated feature. In intensity estimation the method yields estimates with desirable properties and full statements of uncertainty.

KEY WORDS: MCMC; minefield; mixture model; spatial point process; Voronoï tiling.

Contents

List of Tables

List of Figures

1 Introduction

In this paper we consider Bayesian estimation and segmentation for nonhomogeneous Poisson point processes in two dimensions. This work turns out to be easily generalizable to higher dimensions and to have similarities with work done in one dimension.

The principal motivation for the methods considered here can be found in Muise and Smith (1992), Dasgupta and Raftery (1998), Byers and Raftery (1998) and Allard and Fraley (1997). The simplest case is the segmentation of a point process of inhomogeneous rate, for instance the isolation of regions of high density in a point process in the plane, or in a higher dimensional space. Estimation of the rate is also an issue, one that might be intimately bound up with the segmentation as the segmentation methods might serve as a (semi) parametric method of rate estimation.

In Dasgupta and Raftery (1998) the assumption is made that the regions of high density are formed by a mixture of Gaussian distributions on a background Poisson process and model-based clustering (Banfield and Raftery 1993) is used to segment the data. This method is particularly suited to finding very linear features or those approximated by linear forms. In Byers and Raftery (1998) and Allard and Fraley (1997) the point process is assumed to be one of piecewise constant rate, with only two distinct rates being present. The related methods of $k-$th nearest neighbors and Voronoï tile areas, respectively, are used to segment the data.

All of these methods provide some form of classification of the events in the process into high or low density regions. The aim of this paper is to explore methods of gaining more information, such as a region-based estimator with uncertainties and other complex posterior probabilities. The Voronoï tile area based method provides a region but this can be quite crude, and has no uncertainties. Fully Bayesian extensions to the Gaussian cluster approach with solutions via MCMC were considered in Bensmail, Celeux, Raftery, and Robert (1997). We intend to provide the analogous exploration for the approaches in Byers and Raftery (1998) and Allard and Fraley (1997).

Figure 1 shows some point process data that might be segmented into two parts, high density and low density. The second panel shows the data that are estimated to be in the higher intensity region on the basis of their posterior probabilities. Section 5 discusses this example further and compares other methods of segmentation.

The formulation used in the methods examined here is described in Section 2. Section 3 deals with the specifics of intensity estimation while Section 4 deals with aspects of segmen-

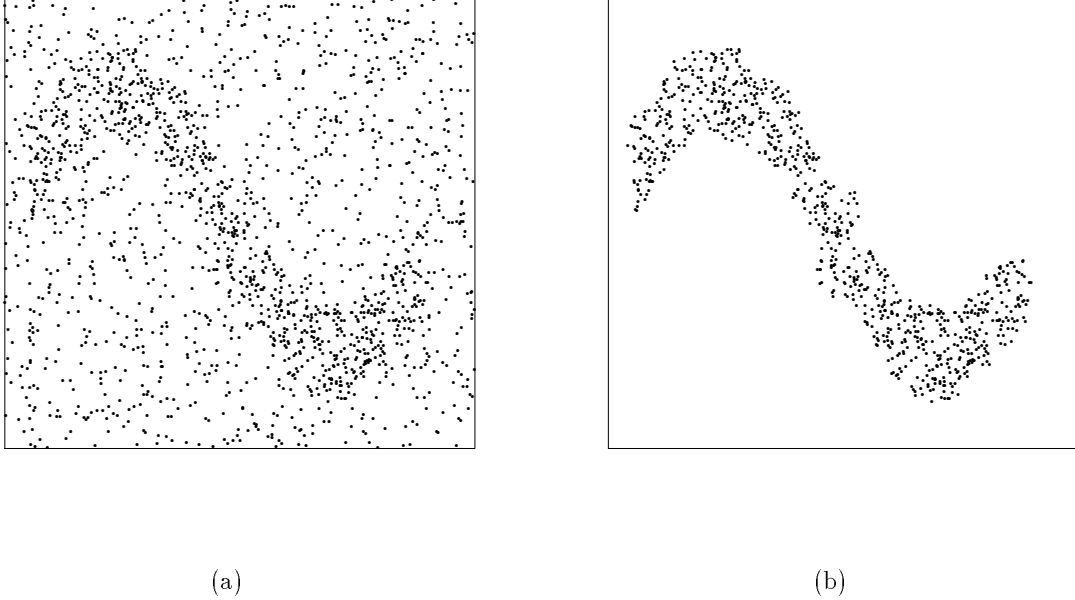(a)   (b)

Figure 1: Left panel shows a point process on a region, right panel shows an estimate of the points in the high rate region of the data. The detection rate is 97.0% while there are 4.2% false positives.

tation. Some examples are shown in Section 5 while Section 6 contains some discussion of this and future work.

2   Proposed Solution Framework

2.1   Formulation

The proposed solution is to approximate a point process as a Poisson point process with piecewise constant rate. Thus the region under consideration, S, will be partitioned up into sets $B_k, k = 1, \cdots, K$ such that $\bigcup_{k=1}^{K} B_k = S$ and $B_i \cap B_j = \emptyset, \forall i \neq j$. The sets will be generated and manipulated by constructing the Voronoï tiling of a set of synthetic generating points, in general unrelated to the data points themselves. Section 2.2 gives a brief account of Voronoï tiling. This framework will be treated in a Bayesian manner and solutions will be sought with Markov chain Monte Carlo methods.

In order to estimate the rate of a process as a function of space, each tile will have its own rate. The posterior distribution of numbers of tiles, tile positions and rates on each tile can

then provide an estimate of the rate and the variability of the estimate as a function of space. A motivation for using this framework is that it has the potential to recover discontinuities in the rate of the process while approximating smooth gradients in intensity.

In order to segment a point process according to intensity the tiles will be grouped into two sets, each set having a different rate. More than two sets could be used in order to allow segmentation into more groups; indeed a variable number of sets could be used.

2.2 Voronoï Tilings

The Voronoï tiling of a set of points $\{(u_k, v_k), k = 1, \cdots, K\}$ in the plane is a partition of the plane into regions $B_k, k = 1, \cdots, K$ such $\forall x \in B_k$, $d(x, (u_k, v_k)) < d(x, (u_l, v_l)), l \neq k$. Voronoï tiles are always convex polygons. An overview of the history of Voronoï tilings with some applications and algorithms for computation can be found in Okabe, Boots, and Sugihara (1992). The partition generated is convenient in that it is easy to find which set, or tile, a point not in the generating set belongs to simply by determining to which generating point it is closest. This partition is also convenient in terms of its specification. Only the generating points need be specified to know the tiling, with the possible addition of a window defining the region under consideration. This framework is similar to that of the simple image segmentation example illustrating reversible jump Markov chain Monte Carlo in Green (1995). There a point process in one dimension has its rate estimated with possibilities for segmentation using a method that might be recast in a Voronoï tiling parameterization. Voronoï tilings extend to higher dimensions and algorithms and code exists to find them in general dimension.

2.3 Markov chain Monte Carlo using Dynamic Voronoï Tilings

In the model we use a piecewise constant rate point process which will be specified and controlled by the Voronoï tiling of a set of synthetic generating points. In the MCMC algorithms used to explore posterior distributions, simple changes in the tiling will enable movement around the state space. Local changes can be made to a tiling by moving, adding or deleting a generating point. Adding and deleting a point induces strictly local changes in terms of neighbor relations, but moving a generating point may theoretically induce larger scale changes, regardless of how small the movement is. On the other hand small movements of generating points can induce much more subtle changes to the tiling than addition or deletion.

3

The task of computing Voronoï tilings we delegate to `deldir`, a package written by Rolf Turner at the University of New Brunswick. The driver routines in RATFOR were used both for our MCMC and via the Splus statistical package for display and interactive investigation purposes.

3   Intensity Estimation

3.1   Formulation

A general Bayesian formulation for the Voronoï tiling model of the intensity of a point process is:

$$
\begin{aligned}
\text{Likelihood} \quad &\propto \quad \prod_{k=1}^{K} \lambda_k^{n_k} \exp\{-\lambda_k a_k\} \\
\text{Priors} \quad & K \sim \text{Poisson}(\nu) \\
& \{\lambda_k; k = 1, \cdots, K\}|K \text{ iid } \sim \Gamma(a, b) \\
& \{ (u_k, v_k); k = 1, \cdots, K\}|K \text{ iid } \sim U\{S\}.
\end{aligned}
$$

Here the a_k are the areas of the Voronoï tiles, or the sets B_k, and the n_k are the numbers of points on each tile. Note that the number of tiles is not fixed in the formulation above, though it could be held at some fixed value. A hyperprior could be placed on the parameter ν, and ν could be updated in the MCMC schedule; this would allow greater flexibility in the ability to tile the region.

The posterior distribution induced by the above formulation is

$$
\pi(K, \{\lambda_k, (u_k, v_k); k = 1, \cdots, K\}|\mathbf{y}) \quad \propto \quad \frac{\nu^K}{K!} \prod_{k=1}^{K} \lambda_k^{n_k + a - 1} \exp\{-\lambda_k(a_k + b)\}.
$$

3.2   MCMC Implementation

In order to fit the models described here we use Markov chain Monte Carlo to simulate from the posterior distribution of the parameters given the data. The algorithms are of the Metropolis-Hastings type, (Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller 1953; Hastings 1970), which have as a special case the Gibbs sampler (Geman and Geman 1984).

4

The samples obtained by these methods from the posterior distributions can then be used to form estimates of the intensity of the process or other quantities of interest.

3.2.1 Fixed Number of Tiles

First we consider the simple case where we use a fixed number of tiles in the partition. We first note the form of the full conditional distributions of the parameters:

$$
\begin{aligned}
\pi(\lambda_k \,|\ldots) &\sim \Gamma(n_k + a, a_k + b), \quad k = 1, \cdots, K \\
\pi((u_k, v_k) \,|\ldots) &\propto \prod_{k=1}^{K} \lambda_k^{n_k + a - 1} \exp\{-\sum_{k=1}^{K} \lambda_k(n_k + b)\}, \quad k = 1, \cdots, K.
\end{aligned}
$$

The $|\cdots$ notation is used to denote conditioning on all other parameters and the data. In the full conditional for (u_k, v_k) the influence of the different values for the parameters is manifested implicitly in the values of a_k and n_k from the tiling induced by changes in (u_k, v_k).

The simplest sampler is one that employs a Gibbs step for each of the λ_k and then proposes moving one or more generating points in a Metropolis-Hastings step. A simple proposal for moving a generating point is uniform in a small square centered at the current position, rendering the move of the Metropolis type.

More general samplers might employ larger scale moves of generating points accompanied by the adjustment of the rate on the tile according to where the proposal generating point sits. It is possible to move more than one, or indeed all of the generating points. The following three schemes were implemented and used together;

- Propose movement of one generating point a small distance.

- Propose movement of a generating point to anywhere in the region. This is more like deleting a generating point and then adding a new one, see the next section. Also the intensity might be proposed to change.

- Propose to move all generating points by very small amounts. This will have the effect of facilitating capture of detail at discontinuities in the intensity.

Whichever of these proposals is tried, the basic elements of the move remain the same: propose movements of points, recompute the new tiling, find the acceptance probability for the new tiling versus the old tiling and update accordingly.

Figure 2 shows two possible methods of altering the tiling by movement or deletion of a generating point. In panel (a) the point to be moved is circled, its new position is shown with an empty circle. The changes to the tiling are shown with the dotted lines. In panel (b) a point to be deleted is circled.

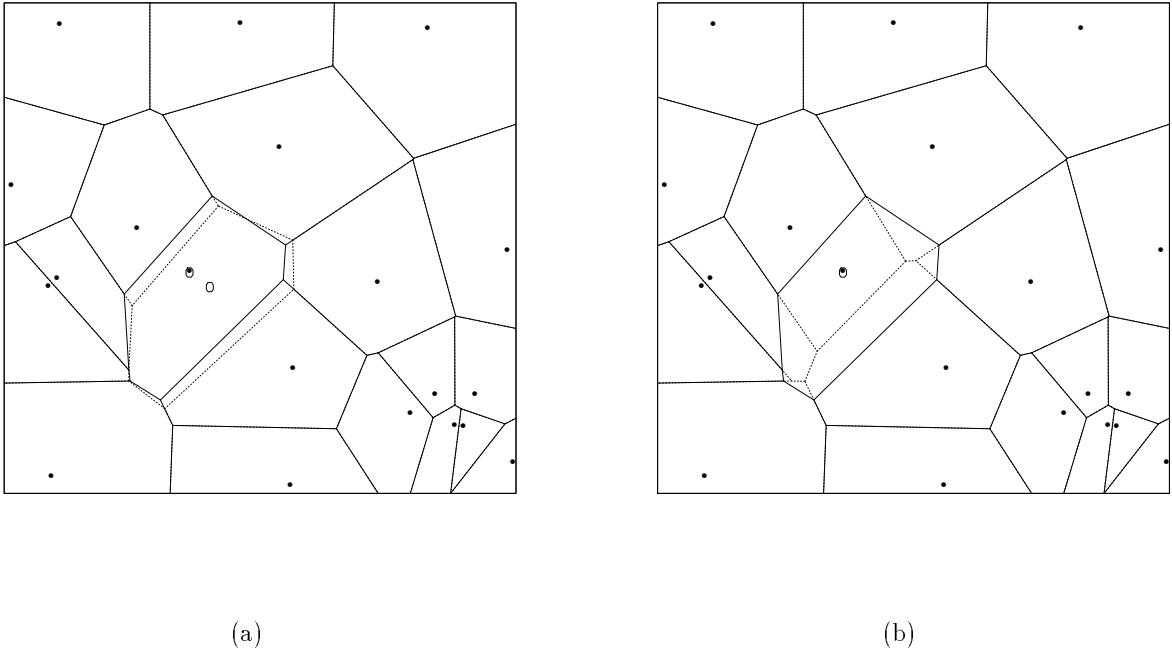<div align="center">(a)   (b)</div>

Figure 2: Examples of two types of proposal for altering the Voronoï tiling in the MCMC algorithm. (a) Proposed movement of a generating point from the ringed point to the empty circle. (b) Deletion (or addition) of the ringed generating point. The dotted lines indicate the induced changes to the tiling.

3.2.2   Variable Number of Tiles

It is desirable to consider a variable number of tiles in the tiling for reasons of flexibility and for mobility of the samplers. For instance, if too small a number of tiles is used, there might not be the ability to capture detail in the process. Regarding mobility, the addition or deletion of a tile can induce large plausible changes in the configuration. In order to sample from a target distribution with a variable number of variables, conventional MCMC must be augmented by reversible jump Markov chain Monte Carlo (Green 1995).

Similar moves will be made to those seen previously, but with the addition of a pair

of moves that propose addition or deletion of a tile, with some suitable adjustment of the intensities on the tiles involved. An addition or a deletion has only local effects on the tiling but even small movements of tile generating points have the capability to change the tiling at long range. This has implications for computational efficiency, and the use of Voronoï tilings in Green (1995) exploits this with the use of the incremental tiling algorithm seen in Green and Sibson (1977).

We make the same type of addition and deletion steps as in the Voronoï example of Green (1995); the proposal distributions and acceptance probabilities are given there. Briefly, the position of a new generating point is sampled from the prior and the rate associated with the induced tile is proposed as $\lambda_{k+1}^* = \tilde{\lambda}_{k+1} v$ where $\tilde{\lambda}_{k+1}$ is the weighted geometric mean of the neighboring tiles and v is drawn from the distribution with density $f(v) = 5v^4/(1+v^5)^2$. The deletion move is the reverse of this construction.

4   Intensity Segmentation

The intention in this section is to segment point processes in the plane based on their intensities. Of particular interest is the case where a process is assumed to have two intensities. Thus the process has piecewise constant intensity, the region under consideration being partitioned into two disjoint sets, each having an intensity. The two sets need not form connected regions.

We consider the case of two regions, one of intensity λ_0, the other of intensity λ_1. The object is to partition the region of interest into two regions and make statements about the uncertainty with which this is done and the estimates of λ_0 and λ_1.

4.1   Formulation

Here S is the region under consideration, A_0 is the area of the sub-region assigned rate μ_0, in which N_0 data points fall, while A_1 is the area of the sub-region with N_1 data points which has rate $\mu_0 + \mu_1$. The rates are parameterized in this manner to give identifiability of the subregions; otherwise one might get swapping of the regions and the rates. In the minefield context these are a noise rate and a mine rate giving an intensity inside a minefield of $\mu_0 + \mu_1$ and outside of μ_0. The classification of the tiles will be coded in the variables z_k, indicating high or low rate. The positions of the tile generating points are $\{(u_k, v_k) : k = 1, \cdots, K\}$. The following formulation results from the above considerations:

$$\text{Likelihood} \quad \propto \quad \mu_0^{N_0}(\mu_0 + \mu_1)^{N_1} \exp\{\mu_0 A_0 + (\mu_0 + \mu_1)A_1\}$$

$$\text{Priors} \qquad \mu_0, \mu_1 \sim \Gamma(a, b)$$

$$K \sim \text{Poisson}(\nu)$$

$$\{z_k; k = 1, \cdots, K\}|K \sim \text{Bernoulli}(p)$$

$$\{ (u_k, v_k); k = 1, \cdots, K\}|K \sim U\{S\}.$$

4.1.1 Fixed Number of Tiles

Here, as in the estimation context, it is possible to use a restricted set of models and fix the number of tiles used in tiling the region. In this case one sweep of the MCMC algorithm might look like the following:

- update the two rates of the process,

- change the classification of a tile from high to low, or vice versa,

- move a generating point, or even more than one.

The posterior distribution is:

$$
\begin{aligned}
\pi(\mu_0, \mu_1, \{z_k, (u_k, v_k); k = 1, \cdots, K\}|\mathbf{y}) \quad \propto \quad & \mu_0^{N_0+a-1}\mu_1^{a-1}(\mu_0 + \mu_1)^{N_1} \\
& \times \exp\{\mu_0(A_0 + A_1 + b) + \mu_1(A_1 + b)\} \\
& \times \prod_{k=1}^{K} p^{z_k}(1 - p)^{1-z_k}.
\end{aligned}
$$

Thus the relevant full conditional distributions are:

$$
\begin{aligned}
\pi(\mu_0, \mu_1 | \cdots) \quad &\propto \quad \mu_0^{N_0+a-1}\mu_1^{a-1}(\mu_0 + \mu_1)^{N_1} \exp\{\mu_0(A_0 + A_1 + b) + \mu_1(A_1 + b)\} \\
\pi(z_k | \cdots) \quad &\propto \quad p^{z_k}(1 - p)^{1-z_k}\mu_0^{N_0+a-1}\mu_1^{a-1}(\mu_0 + \mu_1)^{N_1} \exp\{\mu_0(A_0 + A_1 + b) + \mu_1(A_1 + b)\} \\
\pi((u_k, v_k) | \cdots) \quad &\propto \quad \mu_0^{N_0+a-1}\mu_1^{a-1}(\mu_0 + \mu_1)^{N_1} \exp\{\mu_0(A_0 + A_1 + b) + \mu_1(A_1 + b)\}.
\end{aligned}
$$

The MCMC steps for the segmentation are similar to those for the estimation algorithm, but additional care needs to be exercised in order to retain mobility. The rates of the process are updated together in a bivariate Hastings step as they may be negatively correlated.

8

They are proposed by multiplying by $\exp(U), U \sim \mathrm{Uniform}(-d, d)$. This gives a simple proposal ratio of the ratio of the new to the old values, leading to an acceptance probability $\alpha = \min\{1, R\}$, where

$$
\begin{aligned}
R(\mu_0, \mu_1; \mu_0', \mu_1') &= \left(\frac{\mu_1' \mu_0'}{\mu_1 \mu_0}\right)^a \left(\frac{\mu_0'}{\mu_0}\right)^{N_0} \left(\frac{\mu_1' + \mu_0'}{\mu_1 + \mu_0}\right)^{N_1} \\
&\quad \times \exp\{-A_0(\mu_0' - \mu_0) - (b + A_1)(\mu_1' + \mu_0' - \mu_1 - \mu_0)\}.
\end{aligned} \tag{1}
$$

The z_k are proposed to their opposite value. This gives the acceptance ratio for $z_k = 0 \to z_k' = 1$ of

$$
R(z_k; z_k') = \left(\frac{\mu_0}{\mu_0 + \mu_1}\right)^{-n_k} \exp(-\mu_1 a_k), \tag{2}
$$

where n_k is the number of points that fall on tile k and a_k is the area of tile k. The reverse move has acceptance ratio equal to the inverse of this.

A generating point's movement is proposed as in the estimation algorithm with the possible addition of a change in classification. The new tiling will define A_0' and A_1', while a new μ_0' and μ_1' may be proposed. Then the acceptance probability for the movement of a tile becomes $\alpha = \min\{1, R\}$, where

$$
\begin{aligned}
R((u_k, v_k); (u_k', v_k')) &= (\mathrm{PR}) \left(\frac{\mu_1' \mu_0'}{\mu_1 \mu_0}\right)^{a-1} \frac{\mu_0'^{N_0'}}{\mu_0^{N_0}} \frac{(\mu_1' + \mu_0')^{N_1'}}{(\mu_1 + \mu_0)^{N_1}} \\
&\quad \times \exp\{-A_0' \mu_0' + A_0 \mu_0 - (b + A_1')(\mu_1' + \mu_0') + (b + A_1')(\mu_1 + \mu_0)\}.
\end{aligned} \tag{3}
$$

Here PR is the proposal ratio for any additional adjustments to the rates or the classifications. The proposal for the position (u_k', v_k') is symmetric so does not appear in the acceptance probability.

4.1.2 Variable Number of Tiles

As in the estimation formulation it is desirable to allow a variable number of tiles in constructing the partition of the region. The extension is similar to previously, it involves the addition of an addition and deletion pair of moves in the reversible jump MCMC framework.

The proposal for the position of a new generating point is uniform on the region S as before. Now the new induced tile must be assigned to either the high or low rate, the simplest choice being to generate this assignment randomly from the prior. The two rates do not have to be changed but some perturbation based on the change in the two areas or a random variable might be used. This leads to an acceptance ratio of the form,

$$
\begin{aligned}
R \quad = \quad & \left(\frac{\mu_0' \mu_1'}{\mu_0 \mu_1}\right)^{a-1} \frac{(\mu_1' + \mu_0')^{N_1'}}{(\mu_1 + \mu_0)^{N_1}} \; \frac{1}{p} \frac{\nu}{k+1} |J| \\
& \times \exp\{-b\mu_0' A_0' + (\mu_0' + \mu_1') A_1 - \mu_0 A_0 + (\mu_0 + \mu_1) A_1\}.
\end{aligned}
$$

Here J is the Jacobian for the transformation that maps μ_0, μ_1 into μ_0', μ_1', z_k. The factor of $1/p$ arises from the random assignment of z_k while the Poisson prior for K yields the next term. In this case it is hard to find a good proposal for addition of a point to the tiling.

5 Examples

5.1 Simulated Examples

5.1.1 A Sine Wave

The data shown in Figure 1(a) are from Byers and Raftery (1998). A region of higher density is bounded by two sine waves. We applied the segmentation method to these data to obtain the results in panel (b). The sinusoidal nature of the edges of the region is captured very well. The detection rate is 97.0%, with 4.1% false positives. This result was obtained with a fixed number of tiles, specifically 100, but we chose that to be more that we thought were needed. The add-delete steps had great difficulty being accepted. Greater detail in forming the proposals might improve the acceptance rate for jumps in the variable number of tiles model for these data, but it seems that most of the problem is inherent.

5.1.2 A linear feature

These data are from Dasgupta and Raftery (1998) and were used there to illustrate the `mclust-em` method. Figure 3(a) shows the data and Figure 3(b) shows a probability map for the minefield with the lightest color being zero and the darkest being one. This can be used to yield a segmentation of the data or information about the region in general.
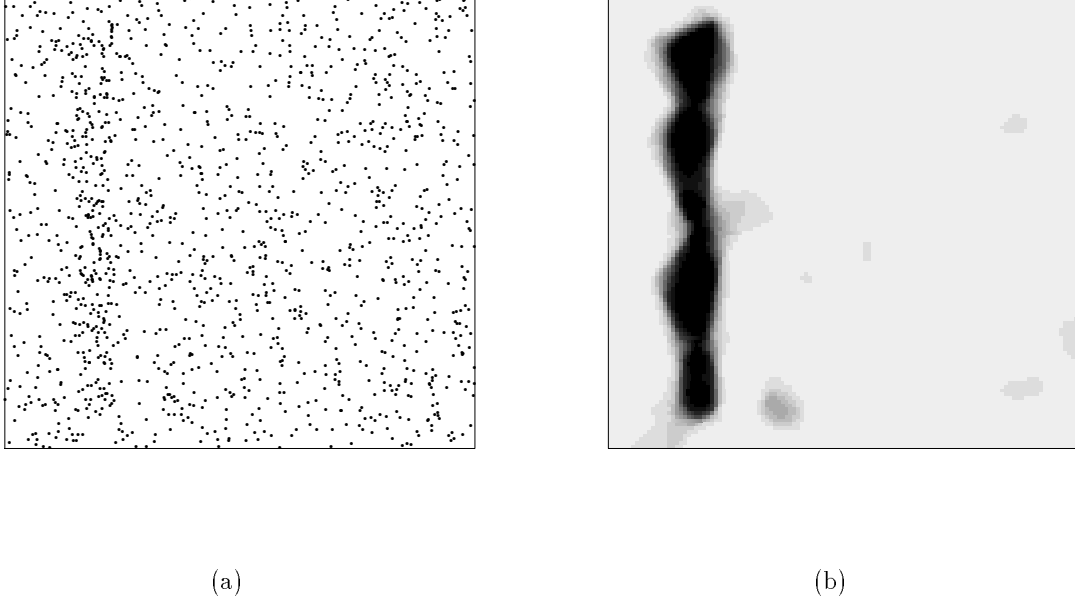
<center>(a)   (b)</center>

Figure 3: (a) Simulated minefield with clutter. (b) Pointwise posterior probabilities of being inside the minefield.

5.2   New Madrid Seismic Region

Data on earthquakes in the New Madrid seismic region are available from the Center for Earthquake Research and Information (CERI) web site at `<http://samwise.ceri.memphis.edu>`. We obtained these data and selected all earthquakes of magnitude greater than 2.5 from 1974 to 1992. To these data we applied the Bayesian intensity estimation and segmentation. The segmentation partitions the data into two regions, one near faults with many earthquakes and one away from major faults with lower, background intensity. The intensity estimation will provide detailed information regarding the intensity of earthquakes as a function of space over the region.

Figure 4(a) shows the earthquake locations. It is clear that the intensity varies in the region, in some places being very low and others quite high. Figure 4(b) shows the pointwise posterior probabilities being in a high-intensity area. In terms of detection of features the results are a linear section joined to a polygonal region with a separate area off on the right of the region.

Figure 5(a) shows the posterior mean estimate of the intensity of occurances in the region. The estimate shows inbuilt locally adaptive properties resulting from the flexible

<center>11</center>

(a)   (b)
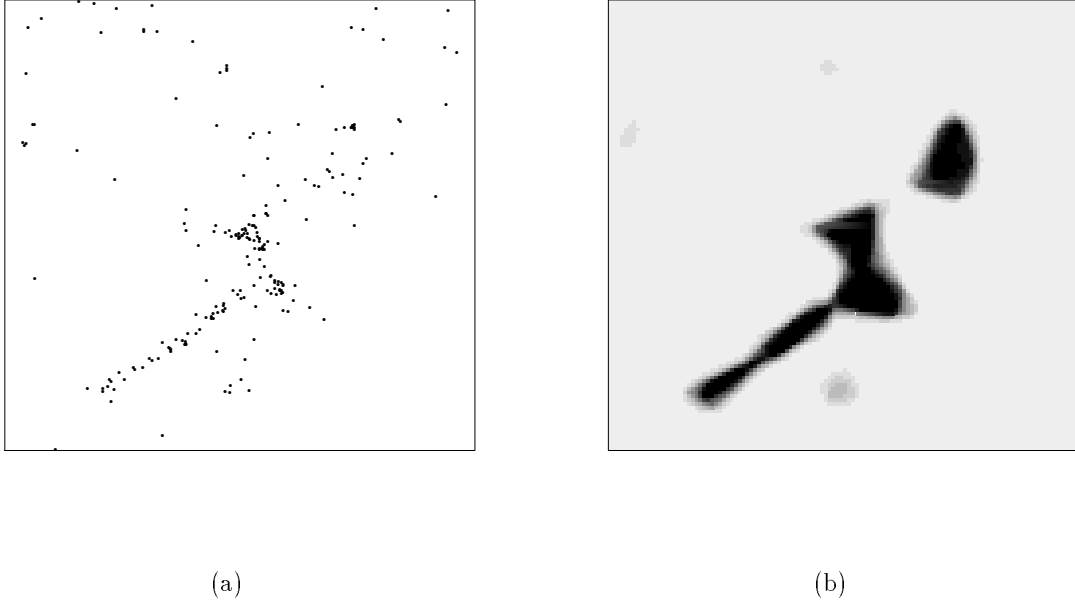
Figure 4: (a) The New Madrid earthquake data, and (b) a Bayesian posterior probability segmentation.

locally constant rate formulation. For example, the high intensity regions are captured with good detail and sharp edges. There is an apparent difference between this and the results of existing non-locally adaptive kernel smoothing methods. For example the above methods implemented in the `Splus` spatial statistics module (Diggle 1983) do not allow for a predominantly smooth estimate that also recovers the very localized regions of high intensity. They also do not provide variability estimates. Figure 5(b) shows the pointwise standard errors of the estimated intensity over the region.

6   Discussion

In this paper we have introduced methods for segmenting and estimating the rate of a spatial point process. They can represent local spatial homogeneity while allowing for discontinuities in the intensity. We use them for Bayesian intensity estimation and segmentation, in both cases yielding information not supplied by other methods.

The segmentation use of the formulation can be regarded as similar in nature to that of Allard and Fraley (1997), but providing a wealth of probabilistic information instead of
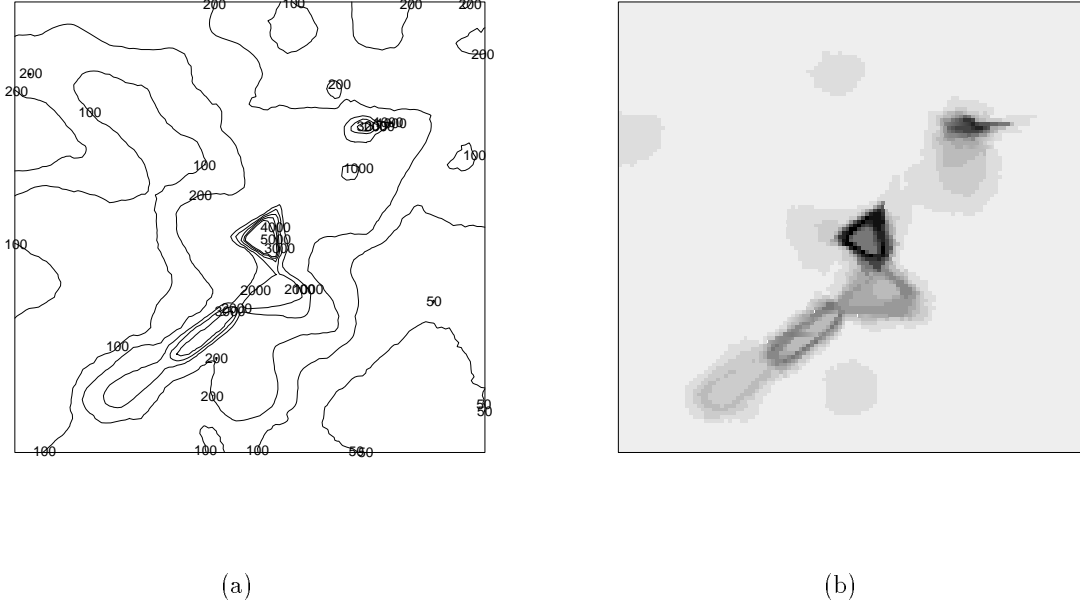
Figure 5: (a) Posterior mean estimate of the intensity of the New Madrid data. (b) Pointwise posterior standard errors.

only classifications. The method of Stanford and Raftery (1997), in its search for purely curvilinear features encounters problems for those in the form of an unstructured region, as evidenced in Figures 13 and 14 of their work. This is also the case in the S* parameterization of `mclust-em` used in Dasgupta and Raftery (1998). The method presented here is a density-based method, as opposed to density and geometry based, and so can capture (curvi)linear or bloblike features together with no additional specification.

Other methods have been proposed that address some aspects of the problems we are trying to solve here. The intensity of a spatial nonhomogeneous point process may be estimated by kernel or other nonparametric density estimation methods (Diggle 1985; Scott 1992). Our proposal, which is also nonparametric, goes beyond this in that it provides formal probabilistic statements of uncertainty about the intensity, and provides an explicit model for segmentation and feature detection. Parametric models have also been used for intensity estimation in nonhomogeneous Poisson processes (Ogata and Katsura 1988); our approach here makes no parametric assumptions about the functional form of the intensity.

Often point process data come with some form of additional information in the form of marks. The formulation we are examining here is easily extendible to allow marks on the

events in the process by use of a model for the marks. In the minefield example a multivariate mark or a more simple expression of confidence might be applicable. For example with each data point there might be a measurement $m_i \in (0, 1)$. Real mines might have higher values while the noise points might have lower values. Marks in the minefield region could then be assumed distributed according to a mixture distribution $pf_1(w) + (1 - p)f_2(w)$ while those outside would be just $f_2(w)$. This expresses the assumption that the data points in the minefield region are a mixture of noise and mines. Further work based on this formulation is underway.

References

Allard, D. and C. Fraley (1997). Non-parametric maximum likelihood estimation of features in spatial point process using Voronoi tesselation. *Journal of the American Statistical Association*. To appear.

Banfield, J. D. and A. E. Raftery (1993). Model based Gaussian and Non-Gaussian clustering. *Biometrics 49*, 803–821.

Bensmail, H., G. Celeux, A. Raftery, and C. Robert (1997). Inference in model-based cluster analysis. *Statistics and Computing 7*, 1–10.

Byers, S. D. and A. E. Raftery (1998). Nearest neighbor clutter removal for estimating features in spatial point processes. *Journal of the American Statistical Association*. To appear.

Dasgupta, A. and A. E. Raftery (1998). Detecting features in spatial point processes with clutter via model based clustering. *Journal of the American Statistical Association 93*. To appear.

Diggle, P. J. (1983). *Statistical Analysis of Spatial Point Patterns*. London. Academic Press.

Diggle, P. J. (1985). A kernel method for smoothing point process data. *Applied Statistics 34*, 138–147.

Geman, S. and D. Geman (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-6*.

Green, P. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrica 82*, 711–732.

Green, P. and R. Sibson (1977). Computing Dirichlet tesselations in the plane. *The Computer Journal 21* (2), 168–173.

Hastings, W. K. (1970). Monte carlo sampling methods using Markov chains and their applications. *Biometrika 57*, 97–109.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. Teller, and E. Teller (1953). Equations of state calculations by fast computing machines. *Journal of Chemistry and Physics 21*, 1087–91.

Muise, R. and C. Smith (1992). Nonparametric minefield detection and localization. Technical Report CSS-TM-591-91, Naval Surface Warfare Center, Coastal Systems Station.

Ogata, Y. and K. Katsura (1988). Likelihood analysis of spatial inhomogeneity for marked point patterns. *Annals of the Institute of Statistical Mathematics 40*, 29–39.

Okabe, A., B. Boots, and K. Sugihara (1992). *Spatial Tesselations: Concepts and Applications of Voronoi Diagrams*. Wiley.

Scott, D. (1992). *Multidimensional Density Estimation.* New York: Wiley.

Stanford, D. and A. Raftery (1997). Principal curve clustering with noise. Technical Report 317, Department of Statistics, University of Washington [Available at `<http://www.stat.washington.edu/tech.reports/tr317.ps>`].