1997

Models and inference for clustering of locations of mines and minelike objects

Noel A. Cressie
Iowa State University, ncressie@uow.edu.au

Andrew B. Lawson
University of Abertay Dundee

Follow this and additional works at: <https://ro.uow.edu.au/infopapers>

Part of the [Physical Sciences and Mathematics Commons](#)

Recommended Citation

Cressie, Noel A. and Lawson, Andrew B.: Models and inference for clustering of locations of mines and minelike objects 1997, 519-530.
<https://ro.uow.edu.au/infopapers/2382>

Models and inference for clustering of locations of mines and minelike objects

Abstract

Mines and mine-like objects are distributed throughout an area of interest. Remote sensing of the area from an aircraft yields image data that represent the superposition of electromagnetic emissions from the mines and mine-like objects. In this article we build a hierarchical statistical model for the reconstruction of mine locations given a point pattern of the superposition of mines and mine-like objects. It is shown how inference on the mine locations can be obtained using Markov chain Monte Carlo methods.

Keywords

models, mines, minelike, objects, locations, clustering, inference

Disciplines

Physical Sciences and Mathematics

Publication Details

Cressie, N. A. & Lawson, A. (1997). Models and inference for clustering of locations of mines and minelike objects. *Proceedings of SPIE - The International Society for Optical Engineering* 3079, Detection and Remediation Technologies for Mines and Minelike Targets II (pp. 519-530).

Models and inference for clustering of locations of mines and mine-like objects

Noel Cressie^a and Andrew B. Lawson^b

^aDepartment of Statistics, Iowa State University, Ames, IA 50011 USA

^bMathematical Sciences Division, School of Informatics, University of Abertay Dundee, Bell St., Dundee DD1 1HG UK

ABSTRACT

Mines and mine-like objects are distributed throughout an area of interest. Remote sensing of the area from an aircraft yields image data that represent the superposition of electromagnetic emissions from the mines and mine-like objects. In this article we build a hierarchical statistical model for the reconstruction of mine locations given a point pattern of the superposition of mines and mine-like objects. It is shown how inference on the mine locations can be obtained using Markov chain Monte Carlo methods.

Keywords and phrases: Cox point process, data augmentation, Markov chain Monte Carlo, Metropolis-Hastings algorithm, Poisson point process, prior distribution

1. INTRODUCTION

The research presented in this paper is concerned with statistical models and analysis for the detection of surface-laid minefields from remotely sensed images. In the past, a minefield has been modeled statistically as a region of elevated intensity in an inhomogeneous Poisson point process.^{1,2} In what is to follow, we generalize this notion to a hierarchical model involving a clustered spatial point process. Specifically, we posit that the mines in a minefield behave according to a Cox process, defined by clusters, random in number and in location,^{3,4} which is superimposed on another Cox process that models the background “clutter” of mine-like objects (e.g., rocks and metal objects other than mines).

In this paper, we shall assume that the data available are a spatial point pattern made up of the superposition of mines (MI) and mine-like objects (ML) in a finite region $A \subset \mathbb{R}^d$. However, individually, they are not distinguishable. A Bayesian hierarchical model is presented in Section 2; because actual minefield data were not available to us, we have demonstrated our methodology on data simulated from the hierarchical model. Computational issues are presented in Section 3 and the simulated data are analyzed in Section 4. Finally, discussion is given in Section 5.

2. MODELS FOR MINES AND MINE-LIKE OBJECTS

In reality, the presence of a mine (MI) or mine-like object (ML) at a given location $y \in A \subset \mathbb{R}^2$ is determined remotely, from airborne sensing equipment. It is often the case that the resulting images are subsequently processed to determine the *locations* of objects (MIs or MLs, which are not individually distinguishable) in the scene. In this case, the data are a spatial point pattern from which one attempts to classify which are the mines in a minefield $A \subset \mathbb{R}^2$.

The first level of our hierarchical statistical model assumes that, conditional on MI intensity function $\mu_a(\cdot)$ and ML intensity function $\mu_b(\cdot)$, MI and ML locations are distributed independently, according to inhomogeneous Poisson point processes (e.g., [5], p. 650). Assume that the MI and ML Poisson intensity functions can be written as,

$$\mu_a(y) = \lambda_a g(y) \left\{ 1 + \sum_{i=1}^{m_a} h_a(y - x_{ai}) \right\}, \quad (1)$$

Further author information -

N.C. (correspondence): Email: ncressie@iastate.edu; Telephone: 515-294-3441; Fax: 515-294-4040

A.B.L.: Email: a.lawson@tay.ac.uk

$$\mu_b(y) = \lambda_b g(y) \left\{ 1 + \sum_{i=1}^{m_b} h_b(y - x_{bi}) \right\} , \quad (2)$$

respectively, where the components of (1) and (2) may be interpreted as follows: λ_a and λ_b are constant background intensity rates for locations of MIs and MLs, respectively, that are linked through

$$\lambda_a = \rho \lambda_b ; \quad (3)$$

$g(\cdot)$ is a function meant to capture common features of the Poisson intensities, due to such things as the topography (hilly versus flat) and the vegetation cover; $m_a, \mathbf{x}_a \equiv (x_{a1}, \dots, x_{am_a})'$, and $h_a(\cdot)$ are cluster number, cluster centers, and cluster function, respectively, for the MIs; $m_b, \mathbf{x}_b \equiv (x_{b1}, \dots, x_{bm_b})'$, and $h_b(\cdot)$ are similarly defined for MLs. For example,

$$h_a(u) = (1/2\pi\kappa_a) \exp\{-||u||^2/2\kappa_a\} \quad (4)$$

$$h_b(u) = (1/2\pi\kappa_b) \exp\{-||u||^2/2\kappa_b\} , \quad (5)$$

which are the cluster functions that shall be used in this paper.

Clearly, (1) and (2) are random functions if some of their constituents are random. At the second level of our hierarchical statistical model, we assume in this paper that

$$m_a, \mathbf{x}_a \sim \text{Strauss}(\beta_a, \gamma_a, R_a) , \quad (6)$$

$$m_b, \mathbf{x}_b \sim \text{Strauss}(\beta_b, \gamma_b, R_b) , \quad (7)$$

where $\text{Strauss}(\beta, \gamma, R)$ is a Strauss point process with intensity parameter $\beta > 0$, interaction parameter $\gamma \in [0, 1]$, and radius $R \geq 0$ (e.g., [5], p. 675). That is, the joint density of m_a and $\mathbf{x}_a$ is,

$$f(m_a, \mathbf{x}_a) \propto \frac{1}{m_a!} \beta_a^{m_a} \gamma_a^{NP(R_a)} ,$$

where $NP(R_a) \equiv \#$ unordered pairs of points within distance R_a of each other. Equivalently, in terms of the unordered locations $\{x_{ai}\}$,

$$f(m_a, \{x_{ai}\}) \propto \beta_a^{m_a} \gamma_a^{NP(R_a)} .$$

An analogous expression is available for $f(m_b, \{x_{bi}\})$.

To be fully Bayesian, one would then put prior distributions on

$$(\rho, \kappa_a, \beta_a, \gamma_a, R_a; \lambda_b, \kappa_b, \beta_b, \gamma_b, R_b) \equiv (\theta_a; \theta_b) . \quad (8)$$

This represents the third level of our hierarchical statistical model. The form of prior distributions used in the example of Section 4 are with fixed Strauss inhibition parameters (γ_a, R_a) and (γ_b, R_b) and fixed $\rho = \lambda_a/\lambda_b$. The (κ_a, β_a) and $(\kappa_b, \beta_b, \lambda_b)$ are all positive parameters sampled from gamma distributions such that zero values are prohibited. The various levels of the hierarchical statistical model used in this paper are specified in (1)–(8). By choosing instead $g(\cdot) \equiv 1$, $m_b = 1$ (a.s.), $h_b(\cdot) \equiv 1$, $\rho \gg 1$, $m_a \in \{0, 1\}$ (a.s.), $h_a(y) = I(y \in B)$, and B a random set such that $B + \{x_{a1}\} \subset A$ (a.s.), we obtain the previously proposed model of a minefield as a region of elevated Poisson intensity.^{1,2}

Return now to (1) and (2). Suppose that $(n_a, \mathbf{y}_a)$ and $(n_b, \mathbf{y}_b)$ are the number and ordered locations of MIs and MLs, respectively. Recall that, conditional on $\mu_a(\cdot)$ and $\mu_b(\cdot)$,

$$n_a, \mathbf{y}_a \equiv (y_{a1}, \dots, y_{a,n_a})' \sim \text{Pois}(\mu_a(\cdot)) , \quad (9)$$

independently of,

$$n_b, \mathbf{y}_b \equiv (y_{b1}, \dots, y_{b,n_b})' \sim \text{Pois}(\mu_b(\cdot)) , \quad (10)$$

where $\text{Pois}(\mu(\cdot))$ denotes an inhomogeneous Poisson point process with intensity $\mu(\cdot)$ (e.g., [5], p. 650). That is, the joint density of n_a and $\mathbf{y}_a$ is,

$$f(n_a, \mathbf{y}_a) = e^{-\mu_a(A)} \prod_{v \in \mathbf{y}_a} \mu_a(v) / n_a! ,$$

where $\mu_a(A) \equiv \int_A \mu_a(u) du$. Equivalently, in terms of the unordered locations $\{y_{ai}\}$,

$$f(n_a, \{y_{ai}\}) = e^{-\mu_a(A)} \prod_{v \in \{y_{ai}\}} \mu_a(v) .$$

An analogous expression is available for $f(n_b, \{y_{bi}\})$ in terms of $\mu_b(\cdot)$. Then the superposition of MIs and MLs leads to the following relations:

$$n = n_a + n_b \quad (11)$$

$$\{y_1, \dots, y_n\} \equiv \left(\bigcup_{i=1}^{n_a} \{y_{ai}\} \right) \bigcup \left(\bigcup_{j=1}^{n_b} \{y_{bj}\} \right) ; \quad (12)$$

notice that the elements of the superposed process, $\{y_i\}$, have their MI/ML designation stripped from them.

Our task now is to write down the joint distribution of all random quantities. A short-hand notation helps enormously in this regard. Suppose $[U]$ denotes the distribution of random variable U and $[U|V]$ denotes the conditional distribution of U given V . Then the joint distribution,

$$\begin{aligned} & [n_a, \{y_{ai}\}; n, \{y_i\}; m_a, \{x_{ai}\}; m_b, \{x_{bi}\}; \theta_a, \theta_b] \\ & \propto [n_a, \{y_{ai}\}; m_a, \{x_{ai}\}; m_b, \{x_{bi}\}; \theta_a, \theta_b | n, \{y_i\}] , \end{aligned} \quad (13)$$

where θ_a, θ_b are given by (8). The goal is to sample from the posterior distribution (13) and to do this we shall use a Markov chain Monte Carlo (MCMC) algorithm. Then, by reporting only the $(n_a, \{y_{ai}\})$ part, a realization from the posterior distribution,

$$[n_a, \{y_{ai}\} | n, \{y_i\}] , \quad (14)$$

is obtained, from which one can make inference on the MIs.

Now, the joint distribution is the product,

$$[n, \{y_i\} | n_a, \{y_{ai}\}, \mu_a(\cdot), \mu_b(\cdot)] \cdot [n_a, \{y_{ai}\} | \mu_a(\cdot), \mu_b(\cdot)] \cdot (6) \cdot (7) \cdot [\theta_a, \theta_b] , \quad (15)$$

where $[\theta_a, \theta_b]$ represents a prior distribution on the various parameters specified in the hierarchical model and θ_a, θ_b are given by (8). Thus, in order to carry out the MCMC algorithm, we need to calculate the first two factors of the product of distributions in (15). From (9) and (10), it is clear that the first two factors are,

$$e^{-\mu_b(A)} \prod_{u \in \mathbf{y} \setminus \mathbf{y}_a} \mu_b(u) \cdot e^{-\mu_a(A)} \prod_{v \in \mathbf{y}_a} \mu_a(v) \cdot I(n \geq n_a) \cdot I(\{y_{ai}\} \subset \{y_i\}) , \quad (16)$$

where $I(C)$ is 1 if the statement C is true and 0 otherwise. Consequently, given $(n, \{y_i\})$, any proposals for $(n_a, \{y_{ai}\})$ must come from the set $0 \leq n_a \leq n$ and $\{y_{ai}\} \subset \{y_i\}$, and realizations from the posterior (13) are eventually obtained after accepting or rejecting proposals for groups of random variables in (13), holding $(n, \{y_i\})$ fixed. Finally then, the posterior probability (13) is proportional to

$$(16) \cdot (6) \cdot (7) \cdot [\theta_a, \theta_b] . \quad (17)$$

Notice that (16) is just

$$[n_a, \{y_{ai}\}; n, \{y_i\} | \mu_a(\cdot), \mu_b(\cdot)] ,$$

which is also equal to the product,

$$\begin{aligned} & [n_a, \{y_{ai}\} | n, \{y_i\}; \mu_a(\cdot), \mu_b(\cdot)] \cdot [n, \{y_i\} | \mu_a(\cdot), \mu_b(\cdot)] \\ & = [n_a, \{y_{ai}\} | n, \{y_i\}; \mu_a(\cdot), \mu_b(\cdot)] \cdot e^{-\mu_a(A) - \mu_b(A)} \prod_{t \in \mathbf{y}} \{\mu_a(t) + \mu_b(t)\} . \end{aligned}$$

Therefore, from this latter expression and (16), we obtain

$$[n_a, \{y_{ai}\} | n, \{y_i\}; \mu_a(\cdot), \mu_b(\cdot)] = \prod_{v \in \{y_{ai}\}} \psi(v) \cdot \prod_{u \in \{y_i\} \setminus \{y_{ai}\}} \{1 - \psi(u)\} \cdot I(0 \leq n_a \leq n) \cdot I(\{y_{ai}\} \subset \{y_i\}), \quad (18)$$

where

$$\begin{aligned} \psi(u) &\equiv \mu_a(u) / \{\mu_a(u) + \mu_b(u)\} \\ &= f(\rho \cdot \{1 + \sum_{j=1}^{m_a} h_a(u - x_{aj})\} / \{1 + \sum_{j=1}^{m_b} h_b(u - x_{bj})\}) , \end{aligned} \quad (19)$$

and $f(w) \equiv w/(1+w)$ is a monotonic increasing function of w .

From (18), it is clear that (conditional on $\mu_a(\cdot)$ and $\mu_b(\cdot)$) the choice of whether an individual location in the observed point pattern is an MI, is a Bernoulli random variable with success probability given by $\psi(\cdot)$, which is a monotonic increasing function of $\mu_a(\cdot)/\mu_b(\cdot)$; notice that the common intensity component $g(\cdot)$ cancels from this calculation. Therefore, if y_1 is the observed location of a MI or ML, $[\mu_a(y_1)/\mu_b(y_1) | n, \{y_i\}]$ is the posterior distribution of the odds ratio that the object is a mine, from which MI/ML classification can be assessed.

3. SIMULATION OF POSTERIOR DISTRIBUTIONS FOR INFERENCE

Using Bayesian methodology, all statistical inferences are based on the posterior distribution given by (13). A closed-form expression for (13) is a nontrivial task, however one can simulate from it using Markov chain Monte Carlo (MCMC) methods. Because we are considering a model with sums of cluster components, where each sum has a random number of terms, it is convenient to use a Metropolis-Hastings (MH) sampler⁶ with the inclusion of reversible jump steps.^{7,8} The basic MH algorithm for sampling cluster processes was developed jointly by A. Baddeley, M. van Lieshout, and A. Lawson.^{9,10,11,12}

For the minefield problem described in Section 2, we observe only the superposition $(n, \{y_i\})$ of MI and ML spatial point processes, $(n_a, \{y_{ai}\})$ and $(n_b, \{y_{bi}\})$, respectively, and hence all other variables in (13) could be viewed as augmented data. Our goal in this paper is to predict certain augmented variables of the MI and ML point processes and ultimately to discriminate between MIs and MLs. Now, the conditional distribution (13) can be generated according to a Markov chain Monte Carlo algorithm that generates successively from the conditional distributions,

$$\begin{aligned} &[n_a, \{y_{ai}\} | n, \{y_i\}, \mu_a(\cdot), \mu_b(\cdot)] , \\ &[m_a, \{x_{ai}\}; m_b, \{x_{bi}\} | n_a, \{y_{ai}\}; n_b, \{y_{bi}\}; \theta_a, \theta_b] , \\ &[\theta_a, \theta_b | n_a, \{y_{ai}\}; n_b, \{y_{bi}\}; m_a, \{x_{ai}\}; m_b, \{x_{bi}\}] . \end{aligned} \quad (20)$$

These conditional distributions are in turn generated according to MH algorithms, as specified below.

3.1 Generation of MI/ML configurations given the superposition

Recall the relationship (18) for the posterior distribution of $(n_a, \{y_{ai}\})$, conditional on knowing the inhomogeneous Poisson intensity functions $\mu_a(\cdot)$ and $\mu_b(\cdot)$. From (20), we can use (18) while conditioning on the current intensity parameters (including any current cluster centers) and, of course, the data. We use a reversible jump method that amounts to a birth-and-death process on the discrete set $\{y_i\}$ of $n = n_a + n_b$ events.

The MH algorithm has two components: The first simulates a *proposal* $(n'_a, \{y'_{ai}\})$, possibly to replace the current value $(n_a^o, \{y_{ai}^o\})$; the second either accepts the proposal or retains the current value with prespecified probabilities (that add to 1). Then, either with the accepted proposal or the retained current value, the cycle is repeated. Because the data $(n, \{y_i\})$ have to be respected by the proposal $(n'_a, \{y'_{ai}\})$, either MIs are added (born) to or taken away (die) from the current value $(n_a^o, \{y_{ai}^o\})$. In our algorithm, this is done one at a time: Given the current value $(n_a^o, \{y_{ai}^o\})$, define the birth probabilities $b(\cdot)$ and death probabilities $d(\cdot)$ according to a uniform distribution,

$$\begin{aligned} b(u) &= (n - n_a^o)^{-1} ; u \in \{y_i\} \setminus \{y_{ai}^o\} , \\ d(v) &= (n_a^o)^{-1} ; v \in \{y_{ai}^o\} . \end{aligned} \quad (21)$$

Then, given the current value $(n_a^o, \{y_{ai}^o\})$, one obtains the proposal $(n_a', \{y_{ai}'\})$ by simulating from the proposal distribution,

$$P((n_a', \{y_{ai}'\}) | (n_a^o, \{y_{ai}^o\})) \equiv \begin{cases} q(n - n_a^o)^{-1} & ; n_a' = n_a^o + 1, \{y_{ai}'\} = \{y_{ai}^o\} \cup u, u \in \{y_i\} \setminus \{y_{ai}^o\} \\ (1 - q)(n_a^o)^{-1} & ; n_a' = n_a^o - 1, \{y_{ai}'\} = \{y_{ai}^o\} \setminus v, v \in \{y_{ai}^o\} \\ 0 & ; \text{elsewhere} , \end{cases} \quad (22)$$

where q is a prespecified probability of a birth being proposed.

The new value is chosen to be,

$$\begin{cases} (n_a', \{y_{ai}'\}), & \text{with probability } \min[1, h] \\ (n_a^o, \{y_{ai}^o\}), & \text{with probability } \max[0, 1 - h] , \end{cases} \quad (23)$$

where

$$\begin{aligned} h &\equiv \frac{P(n_a^o, \{y_{ai}^o\} | n_a', \{y_{ai}'\}) \cdot [n_a', \{y_{ai}'\} | n, \{y_i\}, \mu_a(\cdot), \mu_b(\cdot)]}{P(n_a', \{y_{ai}'\} | n_a^o, \{y_{ai}^o\}) \cdot [n_a^o, \{y_{ai}^o\} | n, \{y_i\}, \mu_a(\cdot), \mu_b(\cdot)]} \\ &= \begin{cases} \frac{(1 - q)}{q} \frac{(n - n_a^o)}{(n_a^o + 1)} \frac{\mu_a(u)}{\mu_b(u)} & ; n_a' = n_a^o + 1, \{y_{ai}'\} = \{y_{ai}^o\} \cup u, u \in \{y_i\} \setminus \{y_{ai}^o\} \\ \frac{q}{(1 - q)} \frac{n_a^o}{(n - n_a^o + 1)} \frac{\mu_b(v)}{\mu_a(v)} & ; n_a' = n_a^o - 1, \{y_{ai}'\} = \{y_{ai}^o\} \setminus v, v \in \{y_{ai}^o\} . \end{cases} \end{aligned} \quad (24)$$

Recall that the new value is obtained after conditioning on the current intensity parameters (and of course the data).

3.2 Generation of MI, ML intensities and their parameters

Generation from $[m_a, \{x_{ai}\}; m_b, \{x_{bi}\} | \theta_a, \theta_b; n_a, \{y_{ai}\}; n_b, \{y_{bi}\}]$ uses relationships (1) and (2), the primary components being the number of clusters (m_a and m_b) and the cluster centers ($\{x_{ai}\}$ and $\{x_{bi}\}$). To obtain a realization, we generate two passes of a birth-death-shifting algorithm, one pass for MIs and one pass for MLs, in a region $D = A \cup \{\text{guard region}\}$, to account for possible cluster centers outside A . Thus we generate a proposal that either replaces the current cluster values or the current values are retained, with probabilities given by a MH algorithm.

The first component of the MH algorithm generates a proposal: We now describe the birth-death-shifting (of cluster centers) algorithm to achieve this. Consider *generic* cluster values $(m, \{x_i\})$, which are cluster number and cluster centers for either MIs or MLs. The current value is $(m^o, \{x_i^o\})$ and the proposal is $(m', \{x_i'\})$. The proposal is obtained from the current value by adding (birth) a cluster center, taking away (death) a cluster center, or shifting a cluster center to a different location. Given the current value $(m^o, \{x_i^o\})$, define these probability distributions by, respectively,

$$\begin{aligned} b(u) &= (m^o)^{-1} \sum_{i=1}^{m^o} k(u - x_i^o) ; u \in D , \\ d(v) &= (m^o)^{-1} & ; v \in \{x_i^o\} , \\ s(u, v) &= (m^o)^{-1} k(u - v) & ; v \in \{x_i^o\}, u \in D , \end{aligned} \quad (25)$$

where, say, $k(x) = \epsilon^{-1} I(0 \leq \|x\| \leq \epsilon)$ and ϵ is a specified bandwidth. Then, given the current value $(m^o, \{x_i^o\})$, one obtains the proposal $(m', \{x_i'\})$ by simulating from the proposal distribution,

$$P((m', \{x_i'\}) | (m^o, \{x_i^o\})) \equiv \begin{cases} (1 - p)q(m^o)^{-1} \sum_{i=1}^{m^o} k(u - x_i^o) & ; m' = m^o + 1, \{x_i'\} = \{x_i^o\} \cup u, u \in D , \\ (1 - p)(1 - q)(m^o)^{-1} & ; m' = m^o - 1, \{x_i'\} = \{x_i^o\} \setminus v, v \in \{x_i^o\} , \\ p(m^o)^{-1} k(u - v) & ; m' = m^o, \{x_i'\} = (\{x_i^o\} \setminus v) \cup u, v \in \{x_i^o\}, u \in D , \\ 0 & ; \text{elsewhere} , \end{cases} \quad (26)$$

where p is a prespecified probability of a shift being proposed and q is a prespecified conditional probability of a birth, given a birth or death is to be proposed.

Now, the MH algorithm for the generation of the second conditional distribution in (20) depends on the ratio,

$$\frac{P((m_a^o, \{x_{ai}^o\})|(m_a', \{x_{ai}'\}))}{P((m_a', \{x_{ai}'\})|(m_a^o, \{x_{ai}^o\}))} \cdot \frac{P((m_b^o, \{x_{bi}^o\})|(m_b', \{x_{bi}'\}))}{P((m_b', \{x_{bi}'\})|(m_b^o, \{x_{bi}^o\}))} \cdot \frac{[m_a', \{x_{ai}'\}; m_b', \{x_{bi}'\}|\theta_a, \theta_b, n_a, \{y_{ai}\}; n_b, \{y_{bi}\}]}{[m_a^o, \{x_{ai}^o\}; m_b^o, \{x_{bi}^o\}|\theta_a, \theta_b, n_a, \{y_{ai}\}; n_b, \{y_{bi}\}]} \quad (27)$$

In (27), each of the first two ratios are, from (26), given by,

$$\begin{cases} \frac{(1-q)}{q} \frac{m^o}{(m^o+1) \sum_{i=1}^{m^o} k(u-x_i^o)} ; m' = m^o + 1, \{x_i'\} = \{x_i^o\} \cup u, u \in D, \\ \frac{q}{(1-q)} \frac{m^o \sum_{x_i^o \neq v} k(v-x_i^o)}{(m^o-1)} ; m' = m^o - 1, \{x_i'\} = \{x_i^o\} \setminus v, v \in \{x_i^o\}, \\ 1 ; m' = m^o, \{x_i'\} = (\{x_i^o\} \setminus v) \cup u, v \in \{x_i^o\}, u \in D, \end{cases} \quad (28)$$

where $m_a^o, \{x_{ai}^o\}$ and $m_b, \{x_{bi}^o\}$ are substituted appropriately into (28).

Now, from Bayes' theorem,

$$\begin{aligned} & [m_a, \{x_{ai}\}; m_b, \{x_{bi}\}|\theta_a, \theta_b; n_a, \{y_{ai}\}; n_b, \{y_{bi}\}] \\ &= \frac{[m_a, \{x_{ai}\}; m_b, \{x_{bi}\}|\theta_a, \theta_b][n_a, \{y_{ai}\}; n_b, \{y_{bi}\}|\mu_a(\cdot), \mu_b(\cdot)]}{[n_a, \{y_{ai}\}; n_b, \{y_{bi}\}|\theta_a, \theta_b]} \\ &\propto \beta_a^{m_a} \gamma_a^{NP(R_a; \{x_{ai}\})} \beta_b^{m_b} \gamma_b^{NP(R_b; \{x_{bi}\})} \cdot (16), \end{aligned} \quad (29)$$

where $NP(R; \{x_i\}) \equiv \#$ pairs of $\{x_i\}$ within distance R of each other. Notice that (29) is simply the product of terms involving MIs with analogous terms involving MLs. Therefore, from (27), $(m_a, \{x_{ai}\})$ is conditionally independent of $(m_b, \{x_{bi}\})$ and the two cluster-center processes can be generated separately. Again, in terms of generic cluster values $(m, \{x_i\})$ we see that, given the current value $(m^o, \{x_i^o\})$, one obtains the proposal $(m', \{x_i'\})$ by simulating from the proposal distribution (26). Then, according to the MH algorithm, the new value is chosen to be,

$$\begin{cases} (m', \{x_i'\}), & \text{with probability } \min[1, h] \\ (m^o, \{x_i^o\}), & \text{with probability } \max[0, 1-h], \end{cases}$$

where

$$\begin{aligned} h &\equiv (28) \cdot [\beta^{m'} \gamma^{NP(R; \{x_i'\})} / \beta^{m^o} \gamma^{NP(R; \{x_i^o\})}] \\ &= (28) \cdot \begin{cases} \beta \gamma^{(\sum_{i=1}^{m^o} I(|u-x_i^o| \leq R))} & ; m' = m^o + 1, \{x_i'\} = \{x_i^o\} \cup u, u \in D, \\ \beta^{-1} \gamma^{-(\sum_{x_i^o \neq v} I(|v-x_i^o| \leq R))} & ; m' = m^o - 1, \{x_i'\} = \{x_i^o\} \setminus v, v \in \{x_i^o\}, \\ \gamma^{(\sum_{i=1}^{m^o} I(|u-x_i^o| \leq R) - \sum_{x_i^o \neq v} I(|v-x_i^o| \leq R))} & ; m' = m^o, \{x_i'\} = (\{x_i^o\} \setminus v) \cup u, v \in \{x_i^o\}, u \in D. \end{cases} \end{aligned} \quad (30)$$

Consequently, we can generate a realization from the second conditional distribution in (20). The third conditional distribution in (20) is that of the intensity parameters θ_a, θ_b , and a realization is likewise generated. Conditional on these two realizations, we go to the first conditional distribution in (20) (Section 3.1), generate a realization, and repeat the cycle. After a burn-in period, all successive realizations come from the stationary distribution of the Markov chain, created so that the distribution is the posterior distribution (13). The next section shows how these realizations can be used to carry out inference.

4. STATISTICAL INFERENCE ON A SIMULATED DATA SET

For purposes of illustration, we have simulated a data set as follows. A process of $m_a + m_b = 10$ cluster centers were distributed uniformly over the unit square. Around each cluster center, four objects were located uniformly in a disk of radius 0.02. Figure 1 displays the simulated realization that will be analyzed in this section. The model given in Section 2 will be fit to these data using the Markov chain Monte Carlo algorithm described in Section 3. (For the data in Figure 1, it is expected that there will be little discrimination between MIs and MLs because there is no distinction built into the simulation.)

Figure 1: Simulated data, simulated from a clustered point process (as described in the text).

The data given in Figure 1 were analyzed via the Markov chain Monte Carlo algorithm given in Section 3. The trio of conditional distributions in (20) were iteratively generated 25,000 times and the Markov chain was assessed for convergence using variance-comparison and distributional-convergence methods.¹³ For example, the last 200 iterations were split into two, namely the penultimate 100 iterations and the final 100 iterations. In Figures 2 and 3, Q-Q plots of n_a (number of mines; see (9)) and κ_b (ML-cluster-function variance; see (5)), respectively, are shown to illustrate convergence of their posterior distributions. (The points on these plots should lie approximately along the line $y = x$ when convergence has occurred.)

Figure 2: Q-Q plot of penultimate 100 iterations versus final 100 iterations from posterior distribution of n_a , the number of MIs.

Figure 3: Same as in Figure 2, except the Q-Q plot shown is for κ_b , the ML-cluster-function variance.

Finally, we are able to recover (an estimate of) the posterior marginal distribution of n_a and $\{y_{a_i}\}$, given the n and $\{y_i\}$ shown in Figure 1. These posterior distributions are estimated by empirical distributions based on the last 1100 iterations of (20). Figure 4 shows the posterior distribution of n_a , the number of mines. Notice how it is peaked about $n/2 = 20$, which is not surprising given that the data were generated indifferently as to which were MIs and which were MLs.

Figure 4: Posterior distribution (estimated): $[n_a|n, \{y_i\}]$.

Figure 5 shows the posterior distribution of $\mu_a(\cdot)$, the Poisson intensity of MIs.

Figure 5: Posterior distribution (estimated): $[\{y_{ai}\}|n, \{y_i\}]$. Shown is a spatially smoothed picture of the proportion of times points of $\{y_i\}$ show up as MIs.

5. DISCUSSION

The simple data set in Section 4 has allowed us to demonstrate that a fully Bayesian approach to minefield detection is possible. The generality of the model allows mines and mine-like objects to be interspersed in irregular patterns, something previous approaches^{1,2} were not able to do. Our model is very general but at the same time can take immediate advantage of local knowledge such as the topography in A (through $g(\cdot)$ in (1) and (2)), of how the MIs are laid (through $h_a(\cdot)$ in (1)), and of the background intensity of MLs (through λ_b in (2)). We have also developed a statistical model that adds another layer on top of the hierarchy should the data come as raw images from remote sensing equipment; this generalization will be reported on elsewhere.

The output from the Markov chain Monte Carlo algorithm is extremely rich. Posterior distributions of any of the augmented variables in (13) can be obtained (through empirical estimates). On the other hand, the algorithm takes a long time to run (the order of a day on a fast workstation); cutting this time down to the order of an hour or two is an important problem that we shall investigate in future research.

ACKNOWLEDGMENT

This research was supported by the Office of Naval Research under grant no. N00014-93-1-0001. The grant also supported a one-month visit by Andrew Lawson to Iowa State University in October, 1996.

REFERENCES

1. R.R. Muise and C.M. Smith, "Nonparametric minefield detection and localization," Technical Memorandum CSS TM 591-91, Coastal Systems Station (Code 2230), Panama City, FL, 1992.
2. A. Dasgupta and A.E. Raftery, "Detecting features in spatial point processes with clutter via model-based clustering," Technical Report No. 195, Department of Statistics, University of Washington, Seattle, WA, 1995.
3. D.R. Cox, "Some statistical methods related with series of events," *Journal of the Royal Statistical Society, Series B* **17**, pp. 129-157, 1955.
4. J. Grandell, *Doubly Stochastic Poisson Processes. Lecture Notes in Mathematics, No. 529*, Springer, New York, 1975.
5. N. Cressie, *Statistics for Spatial Data, revised edition*, Wiley, New York, 1993.
6. W.K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika* **57**, pp. 97-109.
7. C. Geyer and J. Møller, "Simulation procedures and likelihood inference for spatial point processes," *Scandinavian Journal of Statistics* **21**, pp. 359-373, 1994.
8. P.J. Green, "Reversible jump MCMC computation and Bayesian model determination," *Biometrika* **82**, pp. 711-732, 1995.
9. A.B. Lawson, "Discussion on The Meeting on the Gibbs sampler and other Markov chain Monte Carlo methods," *Journal of the Royal Statistical Society, Series B* **55**, pp. 61-62, 1993.
10. A.J. Baddeley and M.N.M. van Lieshout, "Stochastic geometry models in high-level vision," In K. Mardia and G.K. Kanji, eds., *Statistics and Images* **1**, pp. 231-256, Carfax, Abingdon, U.K., 1993.
11. A.B. Lawson, "Markov chain Monte Carlo methods for spatial cluster processes," *Computer Science and Statistics* **27**, pp. 314-319, 1996.
12. A.J. Baddeley, H.Y.W. Cheng, A.B. Lawson, M.N.M. van Lieshout, and N.I. Fisher, "Extrapolating and interpolating spatial patterns," manuscript, 1996.
13. A. Gelman, J.B. Carlin, H.S. Stern, and D.B. Rubin, *Bayesian Data Analysis*, Chapman and Hall, London, U.K., 1995.