J. R. Statist. Soc. B (2007)
69, *Part* 4, *pp.* 701–713

Continuous time modelling of dynamical spatial lattice data observed at sparsely distributed times

Jakob G. Rasmussen and Jesper Møller,

Aalborg University, Denmark

Brian H. Aukema

Canadian Forest Service, Prince George, and University of Northern British Columbia, Prince George, Canada

and Kenneth F. Raffa and Jun Zhu

University of Wisconsin—Madison, USA

[Received August 2006. Revised April 2007]

Summary. We consider statistical and computational aspects of simulation-based Bayesian inference for a spatial–temporal model based on a multivariate point process which is only observed at sparsely distributed times. The point processes are indexed by the sites of a spatial lattice, and they exhibit spatial interaction. For specificity we consider a particular dynamical spatial lattice data set which has previously been analysed by a discrete time model involving unknown normalizing constants. We discuss the advantages and disadvantages of using continuous time processes compared with discrete time processes in the setting of the present paper as well as other spatial–temporal situations.

Keywords: Bark-beetle; Bayesian inference; Forest entomology; Markov chain Monte Carlo methods; Missing data; Multivariate point process; Prediction; Spatial–temporal process

1. Introduction

This paper concerns statistical inference of spatial–temporal processes that are observed on a spatial lattice but only at sparsely distributed time points. Zhu *et al.* (2007) analysed such a data set by using a spatial–temporal autoregressive type of model, which assumes that time is discrete and coincides with the observation times. Here we propose an alternative spatial–temporal model with continuous time, defining a point process for each site of the spatial lattice, where the point processes exhibit spatial interaction; this will be formulated as a spatial–temporal multivariate point process (Daley and Vere-Jones, 2003). Furthermore, we develop Bayesian inference for estimating the model parameters and the times of events. The methodology proposed is illustrated by a subset of the data set that featured in Zhu *et al.* (2007), but we do not attempt to address all the scientific questions there. Instead the focus is on the methodology and our conclusion is that a spatial–temporal multivariate point process model may have several advantages compared with a spatial–temporal autoregressive type of model.

The general type of data that we consider is spatial–temporal data, where we have observed data at sites on a spatial lattice at a discrete set of times. At each time and site we have observed

Address for correspondence: Jesper Møller, Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7 G, 9220 Aalborg Øst, Denmark.
E-mail: jm@math.aau.dk

 1369–7412/07/69701

whether or not an event has occurred. An approach for modelling such data in discrete time is to use a spatial–temporal version of the autologistic model (Besag, 1974). However, the likelihood function for the autologistic model contains an unknown normalizing constant, which makes computation more difficult. Using our alternative approach of a spatial–temporal multivariate point process, the likelihood function is known in closed form, which often will mean that computation is faster for a multivariate point process than for an autologistic model. However, there are several other differences between the two approaches to consider apart from the computational aspects. Very briefly the differences are the inclusion of external information, comparing two data sets observed at different time intervals, consistency, estimation of missing times and prediction. We discuss the differences in detail in Section 6 and point out the various advantages and disadvantages of the two approaches. In this paper we focus on illustrating the differences between the spatial–temporal autologistic model and the spatial–temporal multivariate point process model by using a part of the data set that was modelled in Zhu *et al.* (2007), but the issues that are considered also apply to other discrete and continuous time models, e.g. a spatial–temporal version of the auto-Poisson model compared with a spatial–temporal multivariate point process model.

The paper is organized as follows. Section 2 describes the data which we use to illustrate the methods, and Section 3 introduces the notation needed. Section 4 specifies the spatial–temporal multivariate point process model and prior assumptions, and briefly describes the spatial–temporal autologistic model. Section 5 discusses simulation-based Bayesian inference. Section 6 concludes with a comparison between continuous time and discrete time models for the specific models that are considered in Zhu *et al.* (2007) and the present paper as well as for more general models.

2. Data description

The data set in Zhu *et al.* (2007), which we shall use to illustrate our methods, is from a study of a plantation of red pines near Spring Green, Wisconsin, USA, where the trees were planted on a square grid except that some positions do not have trees (a grey square in Fig. 1 shows the position of a tree). Each tree was examined annually, and three types of data were recorded. The tree's condition (alive or dead) was recorded from 1986 to 1992. The number of *Dendroctonus valens* (LeConte), a bark-beetle which is hereafter called a 'turpentine beetle' that attacks the base of the tree, was recorded from 1987 to 1992. The presence or absence of *Ips* species (predominantly *Ips pini* (Say) and to a lesser extent *Ips grandicollis* (Eichhoff)), another bark-beetle that mass attacks the main stem, was also recorded from 1987 to 1992. The turpentine beetle has one generation per year, with new attacks occurring from late April through June, and each beetle attacking only one tree (Furniss and Carolin, 1980). *Ips* species have two or three generations per year, depending on temperature. They become active in early May and dormant in September. Again, each beetle attacks only one tree. The primary objectives in Zhu *et al.* (2007) were quantifying the relationships between the two types of beetle and the condition of the trees, as well as capturing the spatial and temporal structure of colonization by the two beetle types and the condition of the trees. Of special interest was the fact that a large spatial gap of dead trees appeared.

In the present paper, we model *Ips* species, by using the turpentine beetles as a covariate. Furthermore, we consider only the area immediately around the gap, where Fig. 1 shows the *Ips* species data. Focusing on a subset of the trees leads of course to a loss of important biological information but, since we want to illustrate the methodological aspects, we do not aim at an overly complex model. We specify a spatial–temporal multivariate point process model for *Ips* species, by using a Bayesian setting where we regard what happens between the observation

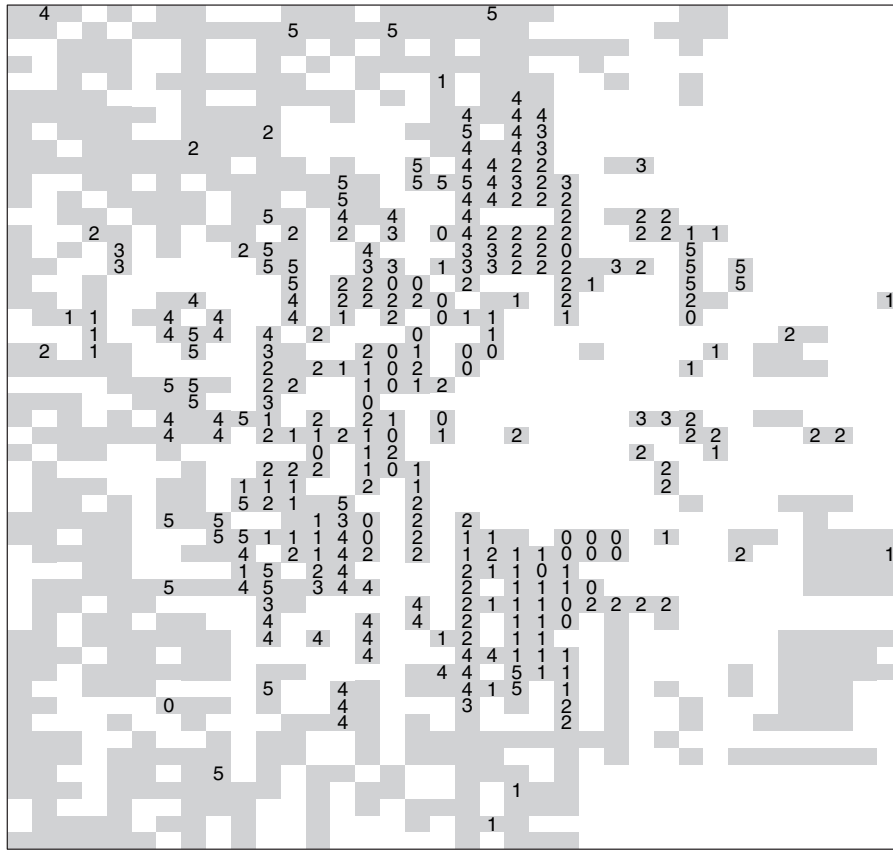

Fig. 1. A white square is a location without a tree or a tree that was already dead in 1986, and a grey square is a tree that was alive in 1986: the numbers indicate which year a tree has been attacked, where 0 corresponds to 1987, 1 corresponds to 1988, etc.; trees without numbers have not been attacked during the period of observation

times as missing data. In short we refer to the spatial–temporal multivariate point process model as a continuous time model and the spatial–temporal autoregressive type of model in Zhu *et al.* (2007) as a discrete time model.

3. Notation

The data were collected during autumn, when *Ips* species were dormant until the beginning of the next period of attack the following spring. For convenience we can therefore assume that the data were observed at the times $k = 0, 1, \ldots, 5$ which correspond to the end of the years $1987, 1988, \ldots, 1992$. Further, we let $k = -1$ correspond to the end of year 1986, and we say that time $t \in \mathbb{R}$ is in year k if $k - 1 < t \leqslant k$.

We let $i = 1, \ldots, 807$ index the sites, where a site corresponds to the location of a tree which is alive at time -1. Let $y_{t,i} = 0$ if site i has not been previously attacked by *Ips* species at time t, $y_{t,i} = 1$ if it has been attacked earlier in the same year and $y_{t,i} = 2$ if it has been attacked in a previous year. For $t \in \mathbb{R}$, we use the notation y_t to denote the vector $(y_{t,1}, \ldots, y_{t,807})$, and the data are $(y_0, y_1, \ldots, y_5)$. By an 'event' at the tree i we mean a transition of the 0–1 process $v_{t,i} = \mathbf{1}[y_{t,i} \geqslant 1]$, where $\mathbf{1}[\cdot]$ denotes the indicator function. We denote the time of this transition

t_i. We do not consider a transition $1 \to 2$ for $y_{t,i}$ as an event, since it is certain that this transition happens at the end of the year at which the event took place. Note that there is a one-to-one correspondence between the process $v_{s,i}$ for $s < t$ (or $s \leqslant t$) and the process $y_{s,i}$ for $s < t$ (or $s \leqslant t$), and we shall mostly consider the process $v_s = (v_{s,1}, \ldots, v_{s,807})$ in what follows.

The process v_s is a particular kind of multivariate point process (or counting process), where each $v_{s,i}$ is restricted to be either 0 or 1. Such a process is specified by the conditional intensity (or hazard) function (Daley and Vere-Jones, 2003): for each tree i, given the history of the process v_s for times $s < t$, let

$$\lambda_{t,i} = \mathbf{E}[\mathrm{d}v_{t,i} | (v_s)_{s<t}]/\mathrm{d}t$$

denote the conditional intensity of the tree being attacked by *Ips* species. In the next section we specify models for the conditional intensity, allowing $\lambda_{t,i}$ to depend on covariate information $(x_{s,i})_{s<t}$, where $x_{t,i}$ denotes the number of turpentine beetles at time t and site i. It will also depend on external information related to *Ips* species activity and on 'neighbouring information'. For a site i, consider its first-, second-, ... order neighbours, which are the (up to) four nearest, four second-nearest, ... sites to i, and let N_i denote the set of first- to fifth-order neighbours of i. Finally, let $u_{t,i} = \mathbf{1}[y_{t,i} = 1]$ and let $n_{t,i} = \Sigma_{j \in N_i} u_{t,j}$ be the number of neighbours of i that are infested with *Ips* species by time t in the same year.

4. Models

4.1. Continuous time model

As usual, $t-$ means the time just before time t. Assume that, for t in year k,

$$\lambda_{t,i} = \mathbf{1}[v_{t-,i} = 0]\, \rho(t)(\psi_0 + \psi_1 n_{t-,i}^{\alpha_1} + \psi_2 n_{(k-1)-,i}^{\alpha_2} + \psi_3 x_{k-,i}) \tag{1}$$

where ρ is a non-negative function, ψ_0, ψ_1, ψ_2 and ψ_3 are non-negative parameters and $\alpha_1 = \alpha_2 = 2$ (this choice and alternative models are discussed at the end of this section). The term $\mathbf{1}[v_{t-,i} = 0]$ is included, since *Ips* species do not attack the same tree twice. The meaning of the other terms in model (1) is given below. Furthermore, given the history v_{t-} and x_{k-} at time t, what happen at different sites are assumed independent.

The function ρ incorporates external information about *Ips* species activity due to seasonal variation. Fig. 2 shows ρ and reflects the fact that *Ips* species has a window of activity of about 4–5 months and normally peak around July. Specifically, for t in year k, $\rho(t) = \varphi\{(t - \mu_k)/\sigma_k\}$ where φ is the standard normal density function and the parameters μ_k and σ_k are determined as follows. Aukema *et al.* (2005) modelled *Ips* species activity as the number of *Ips* species caught in traps every week during the flight period in 2001–2002, by using a linear regression model with various explanatory variables, but only the temperature is available in our study. Therefore we refit the model with temperature as the only explanatory variable, and we estimate μ_k and σ_k by the empirical mean and standard deviation that are obtained from predicting the number of *Ips* species that would have been caught during each week of the kth year. Since $\mu_k - k$ and σ_k do not depend greatly on k, the five normal densities in Fig. 2 look quite similar relative to the years.

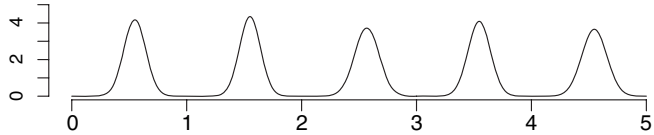

Fig. 2. Function ρ

The function ρ is the same for all sites and is a rough description of *Ips* species activity depending only on the temperature. The term (...) in model (1) adjusts for this. The individual terms in (...) play the following roles. The term ψ_0 is included since the other terms can be 0 and we need then to scale the function ρ. The term $\psi_1 n_{t-,i}^{\alpha_1}$ appears, since *Ips* species may emerge from a tree that was attacked earlier in the year to attack nearby trees. The term $\psi_2 n_{(k-1)-,i}^{\alpha_2}$ appears, since *Ips* species overwinter in the ground close to a previously attacked tree and emerge to attack nearby trees in the following year. The term $\psi_3 x_{k-,i}$ reflects that *Ips* species tend to attack trees that previously have been attacked by turpentine beetles. Since we have only observed which trees the turpentine beetles have attacked at the end of the year, we should ideally also model the turpentine beetles as a continuous time process. However, to keep the model simple we refrain from doing that and instead assume that the turpentine beetles contribute to the conditional intensity throughout the year. Since turpentine beetles usually attack earlier in the year than *Ips* species and $\rho(t)$ is close to being 0 early in the year, this is probably not so unrealistic. Note that, owing to the spatial interaction terms $\psi_1 n_{t-,i}^{\alpha_1}$ and $\psi_2 n_{(k-1)-,i}^{\alpha_2}$, the times of events at neighbouring trees will depend on each other, and such events will not follow the truncated normal distributions that are implied by ρ.

For each year k, the processes $(v_t)_{t<k}$ and $(v_t)_{t>k}$ are conditionally independent given v_k (however, v_t is not a continuous time Markov process). We therefore condition on v_0 and consider the likelihood function based on the remaining data $d = (v_t)_{0<t\leqslant 5}$ (i.e. including the missing data on the time interval [0, 5]). Letting $\psi = (\psi_0, \psi_1, \psi_2, \psi_3)$, the likelihood is (see for example proposition 7.3.III in Daley and Vere-Jones, (2003))

$$L(\psi; d|v_0) = \prod_i \lambda_{t_i,i} \exp\left(- \int_0^5 \lambda_{t,i} \, \mathrm{d}t \right). \tag{2}$$

Furthermore, we specify an improper uniform prior for ψ on $[0, \infty)^4$. Thus the conditional density of ψ given d (and v_0) is proportional to the likelihood (2). A rigorous proof that this conditional distribution of ψ is proper seems difficult, but from a practical point of view we would expect the Markov chain Monte Carlo (MCMC) runs that are described in Section 5.1 to diverge if the distribution was improper. This was not so. Alternatively we could replace the range of ψ by a very large but bounded region.

Actually, before considering model (1), we analysed the model with

$$\lambda_{t,i} = \rho(t) \exp(\psi_0 + \psi_1 n_{t-,i} + \psi_2 n_{(k-1)-,i} + \psi_3 x_{k-,i}),$$

where ψ_0, ψ_1, ψ_2 and ψ_3 are real parameters. This model is of a somewhat similar form to the model in Zhu *et al.* (2007), but a model check along similar lines to those in Section 5.2 showed that the model did not fit the data well. Partly inspired by the form of the Hawkes process (Hawkes, 1971; Daley and Vere-Jones, 2003), we then turned to model (1) but with $\alpha_1 = \alpha_2 = 1$. We observed a misfit and considered therefore alternative values $\alpha_1, \alpha_2 \in \{0, 1, 2\}$. We finally concluded that model (1) with $\alpha_1 = \alpha_2 = 2$ fits best.

4.2. Discrete time model

Since Section 6 compares the spatial–temporal multivariate point process model with the spatial–temporal autologistic model, we shall briefly describe the latter in this section; for more details, see Zhu *et al.* (2007).

In the spatial–temporal autologistic model, we model whether or not tree i has been attacked by *Ips* species at times $k = 0, \ldots, 5$. If tree i has not been attacked before year k, we model the conditional probability $p_{k,i}$ of an attack in year k given the *Ips* species attacks at all other trees

at time k or earlier by

$$\text{logit}(p_{k,i}) = \tilde{\psi}_0 + \tilde{\psi}_1 \tilde{n}_{k,i} + \tilde{\psi}_2 \tilde{n}_{k-1,i} + \tilde{\psi}_3 x_{k,i}, \tag{3}$$

where $\tilde{\psi}_0$, $\tilde{\psi}_1$, $\tilde{\psi}_2$ and $\tilde{\psi}_3$ are real parameters, and $\tilde{n}_{t,i} = \Sigma_{j \in \tilde{N}_i} u_{t,j}$, where $\tilde{N}_i$ denotes the first- and second-order neighbourhood of i. Note that model (3) is allowed to depend on covariates, which in this case is the number of turpentine beetles $x_{k,i}$ that are present in tree i in year k. The four terms in equation (3) have roughly the same interpretation as the four terms in model (1). This is the same model as that used in Zhu *et al.* (2007), except that the terms have been reordered to correspond to model (1), and that we are only modelling *Ips* species colonization (and not tree conditions and turpentine beetle colonization).

Equation (3) is a local specification of a spatial–temporal Markov random field and, using the Hammersley–Clifford theorem (Besag, 1974), we obtain the likelihood function

$$\tilde{L}(\tilde{\psi}; \tilde{d} | v_0) = \frac{1}{c(\tilde{\psi})} \prod_{k=1}^{5} \exp\left\{ \sum_{i: y_{k-1,i}=0} (\tilde{\psi}_0 + \tilde{\psi}_2 \tilde{n}_{k-1,i} + \tilde{\psi}_3 x_{k,i}) + \tilde{\psi}_1 \sum_{i<j: j \in \tilde{N}_i} u_{k,i} u_{k,j} \right\},$$

where $\tilde{d} = (v_t)_{t=1,\ldots,5}$, $\tilde{\psi} = (\tilde{\psi}_0, \tilde{\psi}_1, \tilde{\psi}_2, \tilde{\psi}_3)$ and $c(\tilde{\psi})$ is an unknown normalizing constant. For details on how to treat this normalizing constant as well as other aspects on inference for this model, we refer to Zhu *et al.* (2007).

5. Inference

5.1. Posterior simulation and estimation

The posterior distribution is the conditional joint distribution of ψ and the missing data on [0,5] given the observations $y_0, y_1, \ldots, y_5$. We use a Metropolis-within-Gibbs algorithm to simulate from the posterior distribution—we assume that the reader is familiar with MCMC methods; see for example Robert and Casella (2004).

Specifically, we propose to update each of the four parameters ψ_0, ψ_1, ψ_2 and ψ_3 one at a time by using Metropolis random-walk steps with normal proposal distributions, where the proposal variances are chosen to reach an average acceptance ratio of approximately 0.25 (Roberts *et al.*, 1997). Moreover, within a given year k at a site i either one event $t_i \in (k-1, k]$ has happened, where we do not know the exact value of t_i, or nothing has happened. If an event has happened at site i, we need to simulate the unknown t_i from its conditional distribution given 'everything else', i.e. from the density proportional to $\lambda_{t,i}$, with $t \in (k-1, k]$. We do this by visiting all the sites with events in some predetermined order, updating t_i by an independent Metropolis sampler with proposal density $\rho(t)$, $t \in (k-1, k]$.

Fig. 3 shows plots of the posterior distributions of ψ_0, ψ_1, ψ_2 and ψ_3 based on an MCMC run length of 100000 with a burn-in length of 1000. Note that all the parameters are clearly

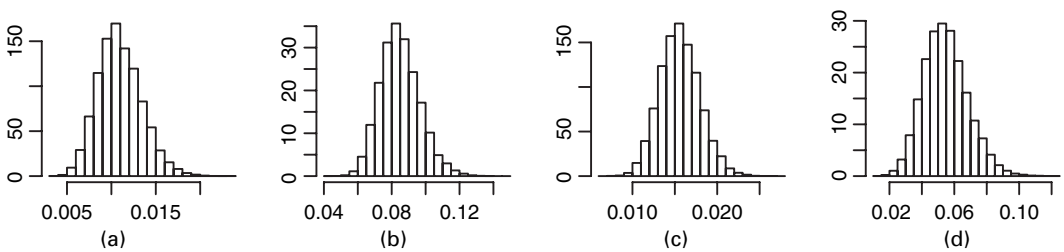

Fig. 3. Posterior distributions of (a) ψ_0, (b) ψ_1, (c) ψ_2 and (d) ψ_3

bounded away from zero. Obviously, since ψ_1 and ψ_2 are significant, the probability that a tree will be attacked increases respectively when more trees in the neighbourhood have been attacked earlier in the year or in the previous year. This leads to several potential biological mechanisms. First, once a beetle has entered a tree and emits aggregation pheromones (Wood, 1982), large numbers of beetles from previously attacked neighbouring trees are available to respond and hence rapidly exhaust the tree's defences (Raffa and Berryman, 1983). Second, factors that predispose trees to being susceptible to attack may be distributed in a highly clustered fashion, and the resident population of beetles again responds to the first entries, thereby generating the pattern observed (Erbilgin and Raffa, 2003). Obviously these are not mutually exclusive. Finally, the significance of ψ_3 in model (1) implies that trees are more susceptible to being attacked by *Ips* species if they have been attacked by turpentine beetles previously.

Fig. 4 shows estimated results for the missing event times t_i. The results are based on an MCMC run length of 100 000 with a burn-in length of 1000 and sampling every 25th missing data set. Figs 4(a)–4(e) show for each year 1–5 the estimated mean value of the t_i, using a grey scale where white means that no event happened at the site and darker values correspond to early events. During the five years the missing data are increasingly further from the initial gap of dead trees, and within each year the mean values tend to be larger at locations further from this gap. Figs 4(f)–4(j) show the standard deviations of the event times, where dark values correspond to small standard deviations. For each year, the standard deviations are about twice as large for isolated attacks as for those occurring in clumps. Since the degree of clumping varies from year to year, the standard deviations in for example year 3 are higher than in the other years. The histograms in Figs 4(k)–4(o) show for each year 1–5 the empirical distribution of all times of events during the year. The histograms are quite similar except for the difference in years, and they look much like the normal densities given by $\rho(t)$ except for a shift about half a month to the right. The posterior distribution contains information about the times of these events, and the shift in the histograms is due to the presence of the term $(\ldots)$ in model (1); see the discussion in Section 4.

5.2. Model checking

Following the idea of posterior predictive model checking (Gelman *et al.*, 1996), we consider posterior predictions: for each $l = 1, 2, \ldots$, we simulate a realization from the posterior distribution and use this when simulating 'new data' $u_{t,i}^{(l)}$ and $v_{t,i}^{(l)}$ from the observation model; see Appendix A. For convenience, let $u_{t,i}^{(0)} = u_{t,i}$ and $v_{t,i}^{(0)} = v_{t,i}$ denote the data.

To check how well the model fits the data, we first consider the number of attacks in each year,

$$w_1^{(l)}(t) = \sum_i u_{t,i}^{(l)},$$

where $t = 1, \ldots, 5$. Fig. 5(a) shows the 2.5-, 50- and 97.5-percentiles estimated from $w_1^{(l)}(t)$ for $l = 0, \ldots, 199$. The values of $w_1^{(0)}(t)$ are also shown in Fig. 5(a). In all years except year 3 where the number of *Ips* species attacks was particularly low, $w_1^{(0)}(t)$ is in the central interval.

Second, let $d(i, j)$ denote Euclidean distance between site i and j. For $\delta \geqslant 1$,

$$w_2^{(l)}(\delta) = \sum_{i<j:d(i,\,j)\in(\delta-1,\,\delta]} v_{5,i}^{(l)} v_{5,j}^{(l)}$$

is the number of pairs of sites between $\delta - 1$ and δ apart and attacked at some time during the observation period. Thus $w_2^{(l)}(\delta)$ quantifies the degree of spatial clustering. Fig. 5(b) shows $w_2^{(l)}(\delta)$ for $\delta = 1, \ldots, 5$ in the same way as in Fig. 5(a). Again there are no discrepancies between the model and the data.

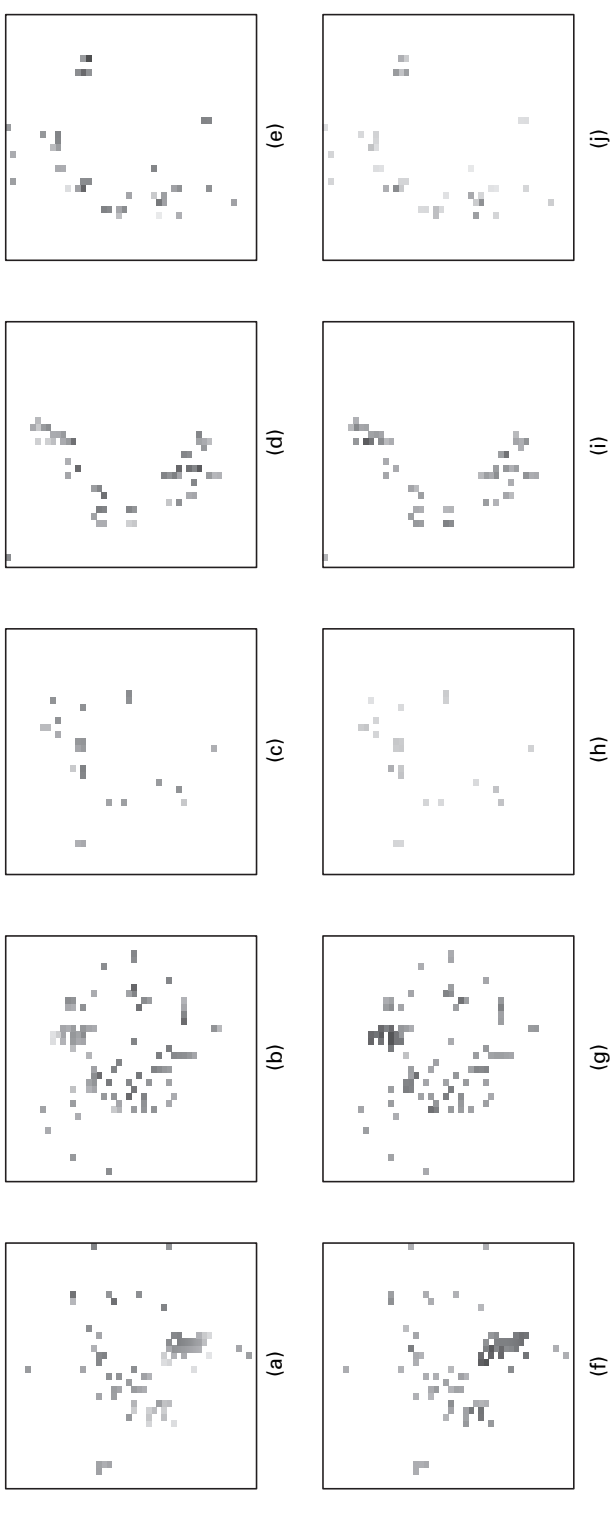

Fig. 4. (a)–(e) Grey scale plots of estimated mean event times, (f)–(j) grey scale plots of estimated standard deviations of event times and (k)–(o) histograms of all event times and $\rho(t)$ within each year: (a), (f), (k), year 1; (b), (g), (l), year 2; (c), (h), (m), year 3; (d), (i), (n), year 4; (e), (j), (o), year 5

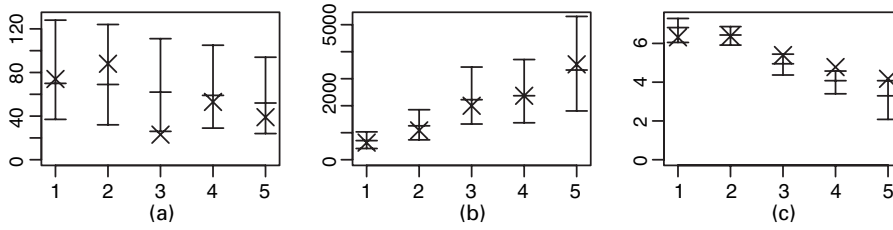

Fig. 5. (a) $w_1(t)$, (b) $w_2(\delta)$ and (c) $\log\{w_3(t)\}$: ×, the data; ‡, 95% central interval and median

Third, to combine temporal and spatial information, consider the number of neighbouring sites that are attacked in years time t apart,

$$w_3^{(l)}(t) = \sum_{i,j,k:\, j \in N_i, k=t,\ldots,5} u_{k,i}^{(l)} u_{k-t,j}^{(l)}$$

for $t = 1, \ldots, 4$, whereas

$$w_3^{(l)}(0) = \sum_{i,j,k:\, j \in N_i, k=0,\ldots,5} u_{k,i}^{(l)} u_{k,j}^{(l)}/2$$

where we divide $w_3^{(l)}(0)$ by 2 to avoid counting all the pairs twice (this is only an issue for $t = 0$). Fig. 5(c) shows $w_3^{(l)}(t)$ for $t = 0, \ldots, 4$ in the same way as the plot for $w_1^{(l)}(t)$, except that we have taken the logarithm of $w_3^{(l)}(t)$ to be able to see what happens at times $t = 3, 4$. For $t = 0, \ldots, 2$, $w_3^{(0)}(t)$ is in the 95% central interval, but for $t = 3, 4$ the model underestimates the number of pairs.

Finally, Figs 6(a)–6(e) show plots of the data for each year $t = 1, \ldots, 5$ where a site i at time t has been coloured black if $v_{t,i}^{(0)} = 1$ and grey otherwise. Simulating 1000 posterior predictions, we let $\tilde{v}_{t,i} = 1$ if $v_{t,i}^{(l)} = 1$ in more than 50% of the simulations and $\tilde{v}_{t,i} = 0$ otherwise. Figs 6(f)–6(j) show plots of $\tilde{v}_{t,i}$. Comparing the two rows in Fig. 6, we see that both the data and the posterior predictive simulations show a clear formation of a large cluster of infested trees in the middle. Furthermore, the clusters seem to be roughly of the same size in the data and the simulations. However, there are some discrepancies between the shape of the cluster in the data and in the simulations: the shape of the cluster is more circular in the data than in the simulations. In spite of this deviation, the general behaviour still seems to have been captured adequately well by the model.

6. Comparison between continuous and discrete time models

We have illustrated how discrete time observations of a spatial–temporal multivariate point process can be analysed by using a Bayesian missing data approach. We conclude with a discussion of the advantages and disadvantages of using continuous time processes compared with discrete time processes in the setting of the present paper as well as other spatial–temporal situations.

6.1. Computation

The most important advantage is the ease of computation. The likelihood function for a spatial–temporal process with discrete time often involves one or more unknown normalizing constants which need to be estimated by using MCMC methods; see for example the Markov random-field models in Besag and Tantrum (2003) and Zhu *et al.* (2007). In contrast the likelihood function for a spatial–temporal multivariate point process is completely specified by modelling the conditional intensity; see equation (2). In Zhu *et al.* (2007), we modelled the data set consisting of the three types of data given by turpentine beetles, *Ips* species and tree conditions. The likelihood

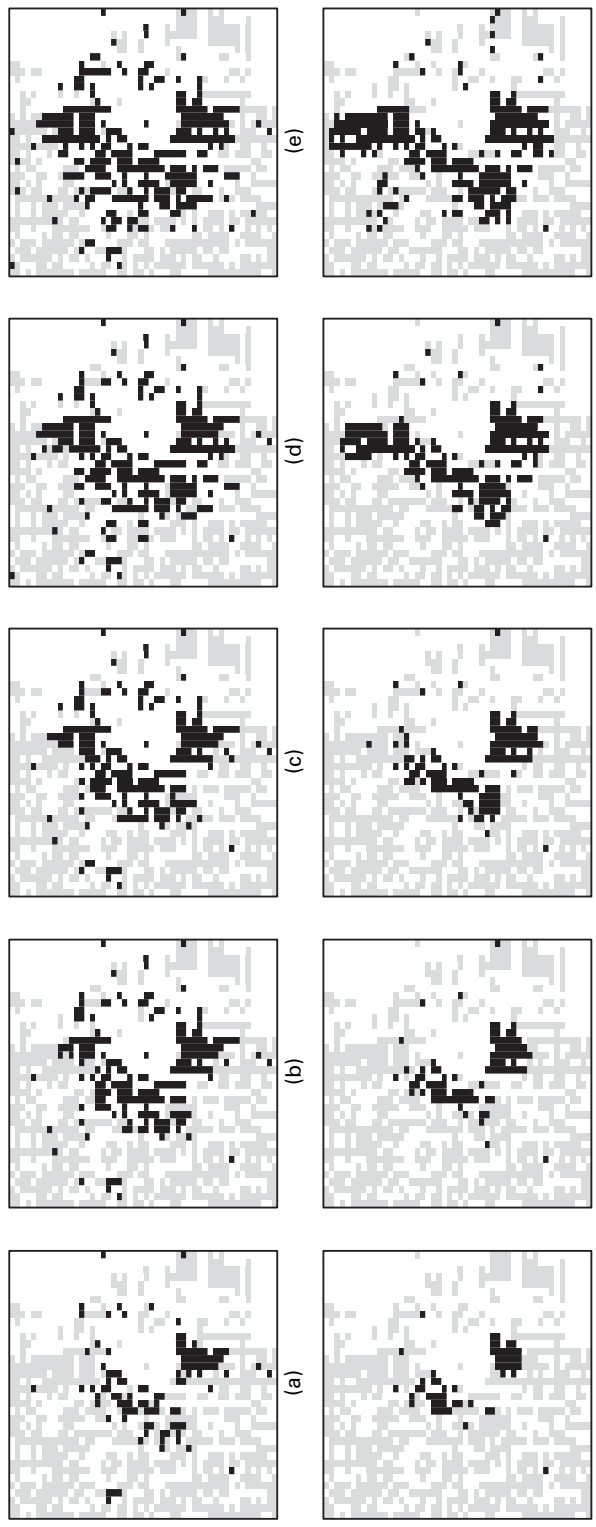

Fig. 6. (a)–(e) A site is coloured black if $v_{t,i} = 1$ and grey if $v_{t,i} = 0$ and (f)–(j) the same as (a)–(e), but for $\bar{v}_{t,i}$: (a), (f) $t = 1$; (b), (g) $t = 2$; (c), (h) $t = 3$; (d), (i) $t = 4$; (e), (j) $t = 5$

function for this model factorized into three terms, one for each type of data, where we can compare the term corresponding to *Ips* species (this corresponds to the discrete time model that was summarized in Section 4.2) and the continuous time model that is used in the present paper, when we include the *Ips* species at all sites (rather than restricting the data to the subset of sites considered so far in the present paper). For this particular comparison, the computation time of the MCMC algorithm for posterior distributions for the discrete time model is roughly 200 times longer than the corresponding computation time for the continuous time model. This improvement in speed means that we have been able to investigate several variations of model (1) (see the end of Section 4.1), which is something that we could not do within practical time limits in the case of the discrete time process for *Ips* species in Zhu *et al.* (2007). However, in Zhu *et al.* (2007) the likelihood terms corresponding to turpentine beetle colonization and tree conditions were easy to specify, and calculations for these terms were very fast. Extending the continuous time model to include turpentine beetles and tree conditions may be much more involved, and a comparison between such a model and the full model in Zhu *et al.* (2007) may turn out differently.

6.2. External information

For a discrete time process as compared with a continuous time process, it may be difficult to incorporate external (or, in a broad sense, covariate) information in the form of another stochastic process observed at a different timescale from that of the discrete time process. For example, the ρ-term in model (1) incorporates external information about *Ips* species activity due to seasonal variation by using measured temperatures. Other similar information, such as soil moisture measured on a daily scale, could be included in a continuous time process in a similar way. Such information can only be incorporated in the discrete time model in Section 4.2 after some form of aggregation. However, we require the external information that was collected in the field across a full season of *Ips* species activity modelled into ρ to obtain a realistic continuous time model for *Ips* species attacks. If we had no such information or only unreliable information was available, the continuous time model approach would be problematic. It would also be a problem if the modelling form of ρ is misspecified. Specifying ρ is time consuming, and this partially offsets the advantage of the shorter computation time for the continuous time model.

6.3. Timescale comparability

Whereas the parameters of two continuous time processes may be compared even if their timescales for observations (i.e. time lengths between observation times) are different, it may not be meaningful to compare the parameters of two discrete time processes with different timescales. The choice of observation times in Zhu *et al.* (2007) is biologically meaningful, since we have annual observations after cessation of insect activity. However, suppose that we had another data set with observations every second year. Further, imagine that we wish to use a model of the same form as in Section 4.2 for both data sets. Then, because of the different timescales, we cannot directly compare the parameters governing for example *Ips* species activity. We could include the missing observation times as missing data, but this would mean that the discrete time process would lose the advantage of not having missing data. Furthermore, if the difference in the timescales is more ill posed, say one data set is measured three times a year and the other is measured eight times a year, this approach may not be practically possible.

6.4. Consistency

Although observed only at discrete times, the types of data that were considered in the present paper come from an underlying continuous time process, and the existence of an underlying continuous time process is not ensured by specifying a discrete time process. Obviously, this is not an issue when we have specified a continuous time model, but our continuous time model only approximates the complexity of the system that is under study. Indeed, *Ips* species colonization does not occur exactly at one time point; it is a complicated process involving hundreds or thousands of beetles attacking over a short period of time. In contrast, the discrete time model reflects the cumulative *Ips* species attacks throughout the past season, and in this sense it is a perfectly sensible model for the system that is under study.

6.5. Estimation of missing data

A continuous time process allows us to model what has happened between observation times, where in our case the event times of within-season *Ips* species attacks can be readily estimated by the MCMC algorithm in Section 5.1. In contrast it is not possible to do this kind of estimation for missing data on the basis of the discrete time model in Section 4.2. The estimates that are based on the continuous time model provide some trustful qualitative results regarding how the times of events depend on the distance to the gap of dead trees and the similar behaviour over the years. However, quantitative results depend much on a careful modelling of external information, particularly the function ρ.

6.6. Predictions

It is straightforward to predict what may happen at any time after the final observation time, applying Ogata's modified thinning algorithm for the continuous time process in the present paper (see Appendix A). The discrete time model in Section 4.2 allows us only to predict what may happen annually. If we are interested only in what happens annually, this can also be obtained from the continuous time model, but note that the shape of ρ still matters even if the exact times of the events are not of interest (this follows from the fact that the process $v_{t,i}$ is a multivariate point process with dependence between the $v_{t,i}$ for $k-1 < t \leqslant k$).

6.7. Conclusion

In conclusion, spatial–temporal processes often suffer from being computationally intensive, and we have demonstrated that, in the present case, using a continuous time process is indeed a viable alternative to a discrete time process.

Acknowledgements

This work was supported by the Danish Natural Science Research Council (grant 272-06-0442, 'Point process modelling and statistical inference'), the US Department of Agriculture 'National research initiative' (2003-3502-13528), the National Science Foundation (DEB0314215), US Department of Agriculture Hatch (WIS01096) and the University of Wisconsin—Madison College of Agricultural and Life Sciences. We are grateful to the Wisconsin Department of Natural Resources for providing the study site.

Appendix A

For model checking in Section 5.2, we need to simulate new data from the observation model. For this we use Ogata's modified thinning algorithm extended to marked processes (Daley and Vere-Jones, 2003; Ogata, 1981).

Factorize the intensity, $\lambda_{t,i} = \lambda_t \; p(i|t)$, where $\lambda_t = \Sigma_i \, \lambda_{t,i}$ is the intensity for the temporal process of events disregarding the location, and $p(i|t) \propto \lambda_{t,i}$ is the probability function for the location given the time of the event. The dependences on the past events (locations and times) of both functions are suppressed in the notation. Further, define the function $l(t)$, which specifies the maximum length that we shall go forward in time in each step of Ogata's modified thinning algorithm, by $l(t) = t + \alpha$ for $k - 1 < t \leqslant \mu_k$, where we let $\alpha = 0.1$, and $l(t) = k$ for $\mu_k < t \leqslant k$. Furthermore, let $m(t) = \max_{s \in [t, l(t)]}(\lambda_s)$. Ogata's modified thinning algorithm is then started at time $t = 0$ and the following steps are repeated until $t > 5$.

(a) Compute $l(t)$ and $m(t)$.
(b) Generate an exponentially distributed variable T with inverse mean $m(t)$ and a uniform variable U on [0, 1] independently of each other.
(c) If $T > l(t)$, set $t = t + l(t)$.
(d) Otherwise, if $t + T > 5$ or $U > \lambda_{t+T}/m(t)$, set $t = t + T$.
(e) Otherwise, let the next event be $t_i = t + T$, where i is generated by using the probability function $p(i|t)$, and set $t = t + T$.

The output is the set of all t_i obtained in step (e).

Combining this algorithm with a sample of the posterior distributions in Section 5.1, we obtain posterior predictions. In practice we obtain a sample of the posterior distributions by taking values from the Markov chains at regular intervals that are chosen such that the sample is effectively independent.

References

Aukema, B. H., Clayton, M. K. and Raffa, K. F. (2005) Modeling flight activity and population dynamics of the pine engraver, *Ips pini*, in the Great Lakes region: effects of weather and predators at short time scales. *Popln Ecol.*, **47**, 61–69.

Besag, J. (1974) Spatial interaction and the statistical analysis of lattice systems (with discussion). *J. R. Statist. Soc.* B, **36**, 192–236.

Besag, J. and Tantrum J. (2003) Likelihood analysis of binary data in space and time. In *Highly Structured Stochastic Systems* (eds P. J. Green, N. L. Hjort and S. Richardson), pp. 289–295. Oxford: Oxford University Press.

Daley, D. J. and Vere-Jones, D. (2003) *An Introduction to the Theory of Point Processes*, vol. I, *Elementary Theory and Methods*, 2nd edn. New York: Springer.

Erbilgin, N. and Raffa, K. F. (2003) Spatial analysis of forest gaps resulting from bark beetle colonization of red pines experiencing belowground herbivory and infection. *For. Ecol. Mangmnt*, **177**, 145–153.

Furniss, R. L. and Carolin, V. M. (1980) Western forest insects. *Miscellaneous Publication 1339*. US Department of Agriculture Forestry Service, Washington DC.

Gelman, A., Meng, X. L. and Stern, H. S. (1996) Posterior predictive assessment of model fitness via realized discrepancies (with discussion). *Statist. Sin.*, **6**, 733–807.

Hawkes, A. G. (1971) Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, **58**, 83–90.

Ogata, Y. (1981) On Lewis' simulation method for point processes. *IEEE Trans. Inform. Theory*, **27**, 23–31.

Raffa, K. F. and Berryman, A. A. (1983) The role of host plant resistance in the colonization behavior and ecology of bark beetles. *Ecol. Monogr.*, **53**, 27–49.

Robert, C. and Casella, G. (2004) *Monte Carlo Statistical Methods*, 2nd edn. New York: Springer.

Roberts, G. O., Gelman, A. and Gilks, W. R. (1997) Weak convergence and optimal scaling of random walk Metropolis algorithms. *Ann. Appl. Probab.*, **7**, 110–120.

Wood, D. L. (1982) The role of pheromones, kairomones, and allomones in the host selection and colonization behavior of bark beetles. *A. Rev. Entomol.*, **27**, 411–446

Zhu, J., Rasmussen, J. G., Møller, J., Aukema, B. H. and Raffa, K. (2007) Spatial-temporal modeling of forest gaps generated by colonization from below- and above-ground beetle species. *J. Am. Statist. Ass.*, to be published.