ESTIMATING LIKELIHOODS FOR SPATIO-TEMPORAL MODELS USING IMPORTANCE SAMPLING

Glenn Marion*
Biomathematics & Statistics Scotland, James Clerk Maxwell Building, The King's Buildings, Mayfield Rd, Edinburgh, EH9 3JZ

Gavin Gibson
Department of Statistics & Actuarial Mathematics, Heriot-Watt University, Edinburgh, EH14 4AS

Eric Renshaw
Department of Statistics & Modelling Science, University of Strathclyde, Glasgow, G1 1XH

This paper describes how importance sampling can be applied to estimate likelihoods for spatio-temporal stochastic models of epidemics in plant populations, where observations consist of the set of diseased individuals at two or more distinct times. Likelihood computation is problematic because of the inherent lack of independence of the status of individuals in the population whenever disease transmission is distance-dependent. The methods of this paper overcome this by partitioning the population into a number of sectors and then attempting to take account of this dependence within each sector, while neglecting that between-sectors. Application to both simulated and real epidemic data sets show that the techniques perform well in comparison with existing approaches. Moreover, the results con rm the validity of likelihood estimates obtained elsewhere using Markov chain Monte Carlo methods.

I. INTRODUCTION

Gibson and Austin (1996) introduced a simple method for estimating parameter likelihoods for a spatio-temporal SI (Susceptible-Infected) model of epidemic spread on a lattice. The techniques were then used to fit the model to observations of Citrus tristeza virus disease spread in an orchard originally reported in Marcus *et al.* (1982). In this model the population members are identified with the vertices of a rectangular lattice, L. Each individual is assumed to be either in state S (susceptible) or in state I (infective). Disease spread is modelled according to the law.

$$\Pr(\sigma(\boldsymbol{x}(t+dt)) = I \mid \sigma(\boldsymbol{x}(t)) = S) \propto \sum_{\{\boldsymbol{y}:\sigma(\boldsymbol{y}(t))\in I\}} F_\alpha(\boldsymbol{y}-\boldsymbol{x})dt. \quad (1)$$

Here $\sigma(\boldsymbol{x}(t))$ denotes the state at time t of the individual at $\boldsymbol{x}$ and the summation is over all infected individuals $\{\boldsymbol{y} : \sigma(\boldsymbol{y}(t)) \in I\}$ at time t. The model formulation assumes that the infective challenges presented to susceptible $\boldsymbol{x}$ by the infectives, $\boldsymbol{y}$, combine additively to determine the overall stochastic rate with which $\boldsymbol{x}$ acquires the disease. F_α is a function, parameterised by α, which describes how the infective challenge presented to $\boldsymbol{x}$ by $\boldsymbol{y}$ is related to the displacement $\boldsymbol{y}-\boldsymbol{x}$. Equation (1) therefore defines the evolution of a stochastic spatio-temporal process. Gibson and Austin (1996) considered the situation where observations, E, are the sets of infected individuals $I_0 \subseteq I_1 \subseteq L$ at two time-points $t_0 < t_1$. They show how a parameter likelihood $L(\alpha \mid E)$ can be defined as

$$L(\alpha \mid E) = \sum_{\omega\in\Omega} \Pr(\omega \mid \alpha), \quad (2)$$

where Ω denotes the set of all orderings of $I_1 \backslash I_0 = \{\boldsymbol{x}_1, ..., \boldsymbol{x}_n\}$, i.e. the set of locations of new infections occurring between the observation times, and $\Pr(\omega \mid \alpha)$ denotes the probability that these locations become infected in the order ω. Note that this approach ignores the exact timings of events which is reflected in the missing constant of proportionality in (1). Gibson (1997b) shows that a simple re-normalisation of time allows this, potentially time-dependent constant to be set to unity. In general, $m \geq 2$ observations $I_0 \subseteq I_1 \subseteq ... \subseteq I_{m-1} \subseteq L$ at times $t_0 < t_1 < ... < t_{m-1}$ may be available. On denoting successive observation pairs at t_{q-1} and t_q by $E_q = \{I_{q-1}, I_q\}$ for $q = 1, ..., m$, the Markovian nature of the system described by (1) means that the likelihood conditional on the set of all observations $\{E_q\}$ may be written in terms of a product of likelihoods based only on observation pairs, whence

$$L(\alpha \mid \{E_q\}) = \prod_{q=1}^{m-1} L(\alpha \mid E_q).$$

In this paper we therefore consider estimation of the likelihood (2) based only on pairs of observations. Computation of $\Pr(\omega \mid \alpha)$ is straightforward (see Gibson and Austin,1996; Gibson 1997b; and the Appendix). However, since $|\Omega| = n!$, it is not in general possible to compute the integral (2) directly, except for small values of n.

Gibson and Austin (1996) use a simple stochastic integration method, simulating orderings ω from a known

*Contact email:glenn@bioss.ac.uk

distribution $g(\omega)$, and appeal to the fact that for ω generated in this way $\Pr(\omega\,|\,\alpha)/g(\omega)$ is an unbiased estimate of $L(\alpha \mid E)$. Using this method with g uniform over Ω they show that it is possible to estimate $L(\alpha\,|\,E)$ for the particular data and models considered, with useful precision. However, subsequent investigations showed that if $F_\alpha(\boldsymbol{y}-\boldsymbol{x})$ defines a predominantly short-range interaction, then strong dependence of $\Pr(\omega\,|\,\alpha)$ on ω can occur. As a result, the variance of $\Pr(\omega \mid \alpha)$ under g is large, and the variance of the estimate $L(\alpha\,|\,E)$ is therefore also large for the uniform sampling model.

Importance sampling (Ripley, 1987) was considered as a means of reducing the variance of the estimates $\Pr(\omega \mid \alpha)/g(\omega)$, by drawing the ω from a known distribution g approximately proportional to $\Pr(\omega \mid \alpha)$. Initial attempts to identify a suitable g with this property were unsuccessful, and subsequent studies (Gibson, 1997a, 1997b; Gottwald et al. 1999) focused on the use of Markov chain Monte Carlo (MCMC, see e.g. Robert and Cassella, 1999; O'Neill and Roberts, 1999) methods to estimate the relative values of $L(\alpha \mid E)$ over the parameter space, by recasting the problem in a Bayesian framework. In this approach, the unknown order ω is treated as a nuisance parameter, and the joint posterior $\pi(\alpha, \omega\,|\,E)$ is investigated using MCMC. When the prior density of α is uniform the marginal posterior $\pi(\alpha \mid E)$ specifies the relative values of $L(\alpha \mid E)$, and $\pi(\alpha, \omega \mid E)$ is proportional to $\Pr(\omega \mid \alpha)$. A variety of algorithms then allow the construction of a Markov chain with limiting distribution $\pi(\alpha, \omega \mid E)$ using the proportionality to $\Pr(\omega\,|\,\alpha)$ in defining the transition probabilities (Gibson, 1997a,b). Once the chain has reached equilibrium, samples $\{(\alpha_i, \omega_i) \mid i = 1, ..., m\}$ may then be used to estimate

$$\pi(\alpha \mid E) = \frac{1}{m}\sum_{i=1}^{m} \pi(\alpha \mid \omega_i, E).$$

If α is constrained to take one of a finite set of values $\{\alpha_l : l = 1, ..., M\}$, then the conditional density

$$\pi(\alpha_l \mid \omega, E) = \Pr(\omega \mid \alpha_l) / \sum_{k=1}^{M} \Pr(\omega \mid \alpha_k).$$

is easily calculated.

In this paper we return to the importance sampling approach. We show that by considering a simplified spatio-temporal process it is possible to identify a distribution $g(\omega)$ which considerably reduces the variance of $\Pr(\omega \mid \alpha)/g(\omega)$ in comparison with the uniform case. There are several benefits to pursuing this approach.

• Alternative integration techniques such as importance sampling are useful for confirming the validity of results obtained by MCMC. The convergence of MCMC methods is the subject of much interest, and in many cases it is not yet possible to determine whether a sequence of samples from a Markov chain is truly representative of the stationary distribution (Cowles and Carlin, 1996). This is true for the results of Gibson (1997a,b) although sample based tests did not indicate convergence problems. This issue may be circumvented by use of perfect simulation (Propp and Wilson, 1996) which generates samples from the equilibrium distribution of the Markov chain. However, practical implementation of this approach requires the identification of an ordering on state space, which is preserved under the evolution of the Markov chain. Typically, identification of such an ordering is not straightforward and we have been unable to do so for the estimation problem considered here.

• To formulate a suitable g one might make qualitative, simplifying approximations regarding the spatio-temporal process. The validity of these should be reflected by the variance of the corresponding likelihood estimator. Therefore, by formulating and testing a distribution g in this way we may be able to gain insight into the essential dynamics of the spatio-temporal process of interest.

• Direct estimates of the likelihood obtained via importance sampling would greatly simplify the process of model comparison described by Gibson and Renshaw (2001) who use MCMC sampling to conduct a generalized likelihood ratio test.

• The methods described here need not be used in isolation from other techniques. For example, g may be used as the proposal distribution for an MCMC sampler, which should improve the mixing rate of the overall algorithm.

In the next section we describe the method of construction of our sampling distribution g which is designed to be approximately proportional to $\Pr(\omega\,|\,\alpha)$. Then in Section 3 we apply the method to simulated data sets in order to quantify the improvement in performance in comparison with the case where g is uniform. Finally, in Section 4 the technique is applied to a real data set previously analysed using MCMC by Gibson (1997a).

II. LIKELIHOOD ESTIMATION USING IMPORTANCE SAMPLING

Because the models considered have a simple Markov structure throughout, it suffices to discuss the computation of likelihoods for the case where an epidemic is observed at two distinct time points, t_0 and t_1. However, data from several sequential pairs of observation times may each be treated as described below, and combined using the Markov property. We retain the notation introduced in Section 1. The approach taken to construct the appropriate sampling distribution g is the following. We decompose the lattice as a partition of k disjoint subsets $L = L_1 \cup ... \cup L_k$. Then we consider

a modified spatio-temporal model for the development of the epidemic between t_0 and t_1. The modified model, denoted M', is as described in the Introduction with the following addendum. We assume that a susceptible at $\boldsymbol{x}$ can acquire the disease:

- either from an individual $\boldsymbol{y} \in I_0$;
- or, from an individual $\boldsymbol{y} \in L_j \cap (I_1 \backslash I_0)$ where $\boldsymbol{x} \in L_j$.

Now for $1 \leq j \leq k$, let the set $X_j = L_j \cap (I_1 \backslash I_0)$, with $|X_j|$ elements. In addition, let ω^j be an ordering of X_j and consider the probability

$$\phi_j(\omega^j) \;=\; \Pr\{\text{the first } |X_j| \text{ infections (after } t_0) \text{ in } L_j \text{ occur at } X_j \text{ in the order } \omega^j \mid M'\}$$

Since for M' the infection process within L_j is independent of that in $L_i, i \neq j$, then $\phi_j(\omega^j)$ depends only on ω^j. Moreover, if $|X_j|$ is sufficiently small then it is possible to evaluate $\phi_j(\omega^j)$ for all possible choices of ω^j.

The algorithm for generating orderings of $I_1 \backslash I_0$ is now described.

1. For each j, $1 \leq j \leq k$, generate an ordering ω^j of X_j from the distribution ϕ_j.

2. Generate a total ordering ω of $I_1 \backslash I_0$ by repeatedly selecting the first previously unselected site in ω^j, for j selected with probability proportional to the number of previously unselected sites in X_j.

The densities ϕ_j are easily computed as shown in the Appendix. Moreover, if $\boldsymbol{J}$ denotes the sequence of $j_1, ..., j_n$ selected in (2) above then the density $f(\boldsymbol{J})$ can be easily computed (see Appendix). Therefore the density of the ordering ω generated by (1) and (2) above can be expressed as

$$g(\omega) = f(\boldsymbol{J}) \prod_{j=1}^{k} \phi_j(\omega^j).$$

By partitioning the lattice in this way we attempt to take account of chains of causality within each sector L_j and the effect that this may have on the probabilities of different temporal orderings in the sector, while neglecting the effect of chains of causality (infection) which cross between sectors. So the method shares something of the philosophy of cluster approximations (e.g. Filipe and Gibson, 1998) to spatio-temporal model behaviour in that it neglects aspects of interactions between clusters. When $k = 1$, the density $g(\omega)$ is precisely proportional to $\Pr(\omega \mid \alpha)$ and the original problem is recovered. By selecting the L_j so that the number of new infections occurring in each is manageable (so that the densities ϕ_j can be computed and stored) the generation of ω and the computation of $g(\omega)$ become computationally feasible. The simplification of the dynamics of spatio-temporal lattice processes by partitioning has also been considered by Rand and Wilson (1995).

III. RESULTS WITH SIMULATED EPIDEMICS

In this and the following section we apply our importance sampling procedure to analyse first simulated and then real epidemic data sets, and quantify the improvements over the case where the orderings ω are sampled uniformly. First we explore the application to data generated by a model infection process of the type described by equation (1) where the force of infection is described by a power-law, namely

$$F_\alpha(\boldsymbol{y} - \boldsymbol{x}) = \mid \boldsymbol{y} - \boldsymbol{x} \mid^{-2\alpha} . \tag{3}$$

The larger the value of α the shorter the range over which infection is likely. We consider this infection process on a lattice of 20×20 sites with unit spacing. In order to generate test data, initial infections on the lattice are drawn randomly from a multinomial distribution with infection probability 0.1; in the realization used here 46 primary cases are created which corresponds to a relatively severe initial infection. Approximately 50 secondary infections are then generated by the process (1) with the power-law infection force (3). This is performed for two power-law exponents $\alpha = 2$ and $\alpha = 2.5$. Figures 1a and 2a show that the results obtained by applying random sampling to estimate the likelihood $L(\alpha \mid E)$ from both sets of data are rather poor. This is because the infection process is dominated by predominantly short-range interactions and $\Pr(\omega \mid \alpha)$ is strongly dependent on the ordering ω of secondary infections. This situation is one in which our importance sampling technique should perform well. However, before it can be applied the lattice L must be decomposed into appropriate disjoint subsets L_j, $j = 1, ..., k$. Figures 1d and 2d show the primary and secondary infections for each data set, and the regions used. The L_j are chosen subject to some upper limit on the number of secondary cases within a given region, such that they contain a spatially distinct group of secondary infections. Under the action of infection process (1) these regions are likely to contain causally linked chains of secondary infections, and importance sampling based on such regions should result in more precise estimates of the likelihood compared with those obtained from random sampling.

In the examples considered here only rectangular regions are considered, although more general geometries could be used. Furthermore, for reasons of computational effort the number of secondary infections per region is limited to a maximum of nine. For the data set generated with $\alpha = 2$ Figure 1b shows the estimates of $L(\alpha \mid E)$ obtained by applying the importance sampling method with regions (shown in Figure 1d) containing less than 7 secondary infections. Comparison with Figure 1a shows a clear improvement over random sampling with a marked reduction in the variance of the likelihood estimates. Importance sampling (results not shown) with regions containing less than 5 secondary infections show a similar improvement over random sampling to those of Figure 1b with a corresponding reduction in computa-

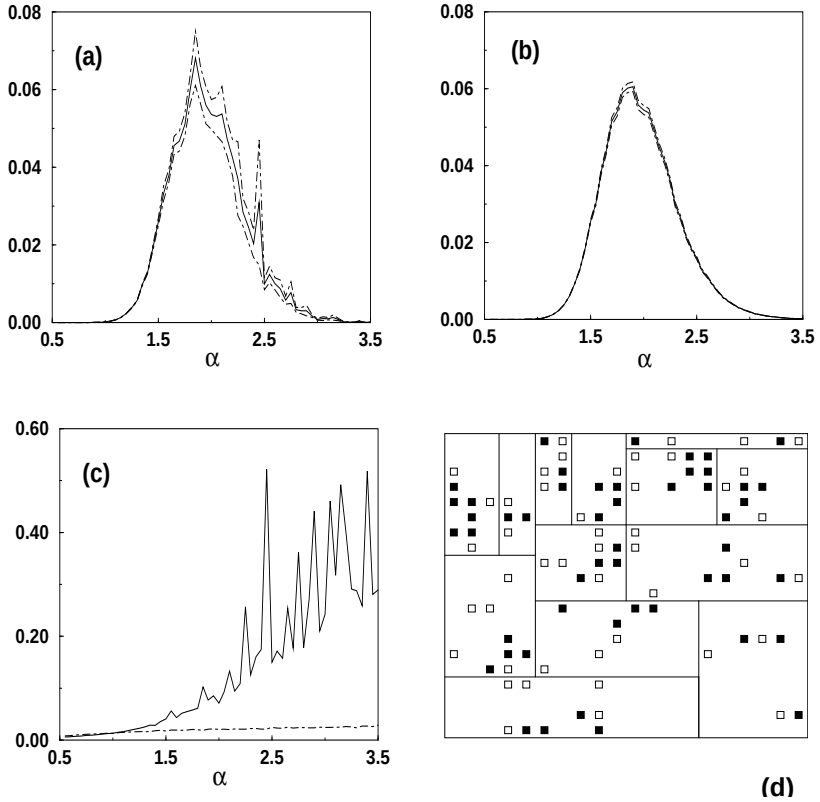

FIG. 1: Parameter likelihoods estimated from data generated by the infection process (1) with a power-law contact rate (3) with α = 2.0. (a) shows the estimate of $\mathsf{L}(\alpha \mid E)$ (solid curve) and $\pm$ two standard deviation interval (dot-dashed curves) obtained by random sampling, whilst (b) shows that obtained by using the importance sampling procedure (see text for details). (c) shows the coe cient of variance of the estimates of $\mathsf{L}(\alpha \mid E)$ obtained for random (solid curve) and importance (dot-dashed curves) sampling. In (d) the open squares show the relative positions of the primary infections, the solid squares those of the secondary infections, and the lines delineate the regions used in the importance sampling algorithm.

tional effort. This reduction of variance is further illustrated by the coefficient of variation shown in Figure 1c where the improvement gained by using our method of importance sampling increases markedly as the range of the interaction reduces (α increases).

Figures 2b and 2c show the application of the importance sampling procedure to data generated by the power-law infection process with $\alpha = 2.5$ starting with the same 46 primary infections. In this case regions containing less than 9 secondary infections (Figure 2d) are used and, as anticipated, the improvement over random sampling is even more pronounced in this case where the interactions are shorter in range.

In order to assess the impact of region choice on the performance of the importance sampling algorithm three further designs are explored for the case $\alpha = 2.5$. Figure 3c shows the first of these, which is intended as an example of extremely poor region design; the boundaries have been chosen so as to break up clusters of secondary infections to the extent that no adjacent pairs of secondary infections appear in the same region. Figures 3a and 3b show that, as expected, the precision of the estimates is no better than for random sampling. Figure 4c shows a slightly improved design in which regions enclose at most adjacent pairs of secondary infections. However, since the data contain clusters of three or more adjacent pairs this design is also expected to perform poorly. This result is confirmed by Figure 4a which shows little improvement over random sampling. Moreover, Figure 4b shows that these results are poor in comparison to those obtained from importance sampling based on the larger regions shown in Figure 2d. Figure 5c shows a similar design in which regions are slightly enlarged, but remain smaller than those of Figure 2d, and enclose at most three adjacent secondary infections. Figures 5a and 5b show that importance sampling based on this design produces estimates of the likelihood whose precision is a considerable improvement over random sampling, but still poor compared to the results of Figure 2. The lesson is clear: regions should be chosen to enclose clusters of adjacent secondary infections which are as large as possible, but the number of secondary infections within a region must be constrained by the need to keep the explosion of possible orderings of these to a manageable level. However, our results suggest that useful precision may be obtained even when regions are forced to break-up natural clusters

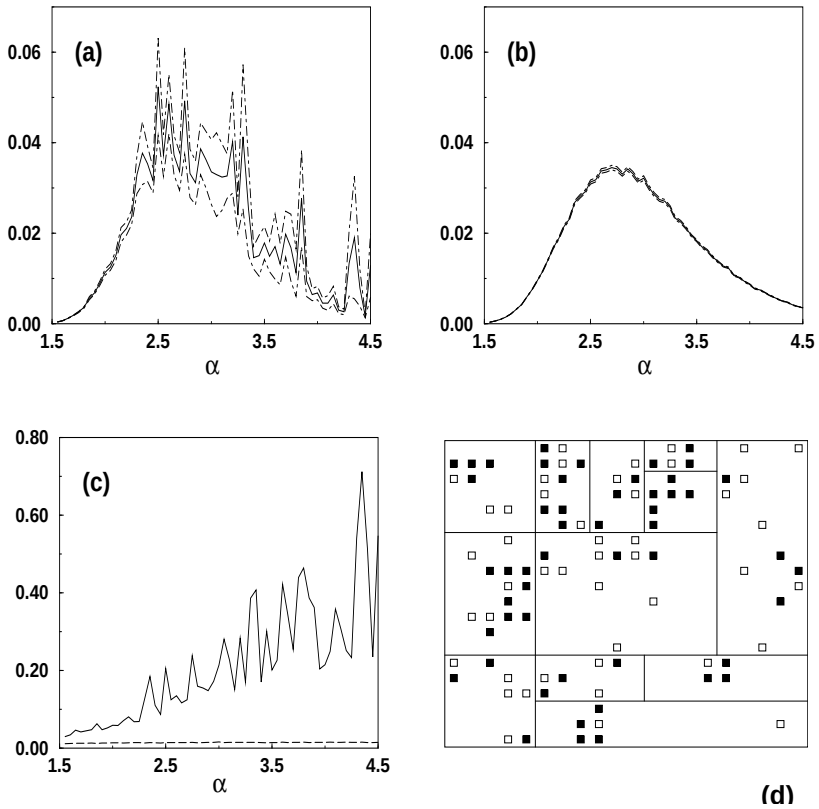

FIG. 2: As Figure 1, but with data generated by infection process (1) with a power-law contact rate (3) with α = 2.5.

of secondary infections, so long as they enclose reasonable subsets of such clusters.

IV. APPLICATION TO CITRUS TRISTEZA VIRUS DATA

Here we apply importance sampling to the Citrus tristeza virus data set studied by Gibson (1997a). Figure 6d shows the positions of the observed primary and secondary infections, along with the regions (each containing less than 8 secondary infectives) used in the importance sampling procedure. Gibson (1997a) uses MCMC to estimate model parameters in a generalisation of (1), namely

$$\begin{aligned}\Pr(\sigma(\boldsymbol{x}(t+dt)) = I | \sigma(\boldsymbol{x}(t)) = S) \\ \propto \beta dt + \sum_{\{\boldsymbol{y}:\sigma(\boldsymbol{y}(t))\in I\}} F_\alpha(\boldsymbol{y}-\boldsymbol{x})dt \end{aligned} \quad (4)$$

where β is the background infection rate and the contact rate is the power-law (3). Gibson (1997a) estimates the joint likelihood $L(\alpha, \beta \mid E)$, and the maximum likelihood parameter estimates found within $\alpha \in [0, 3.5]$ and $\beta \in [0, 1.12]$ are $\alpha = 3.5$ and $\beta = 0.833$. Although these estimates are on the boundary of the parameter space considered, the likelihood is relatively flat for large values of α (> 3.5) which corresponds to a quasi-nearest-neighbour interaction. Figure 6a shows the estimate of the conditional likelihood $L(\alpha \mid E, \beta = 0.1)$ obtained using random sampling, whilst Figure 6b shows the corresponding estimates obtained from importance sampling (based on the regions shown in Figure 6d). The resulting variance reduction is quantified by the coefficient of variation (Figure 6c). The value of the background infection rate is chosen to be less than the maximum likelihood estimate as in this region of parameter space the process of infection predominates. Moreover, $L(\alpha \mid E, \beta = 0.1)$ has a well-defined maximum in contrast to the profile likelihood $L(\alpha \mid E, \beta = 0.833)$ which, as noted above, levels off as α increases. The results show a striking improvement in the precision obtained from importance sampling compared with that obtained from random sampling. Although the results are not shown, the importance sampling method was applied to a second Citrus tristeza virus data set, and the resulting estimates of the likelihood agree well with those obtained by Gibson (1997b) using MCMC.

V. CONCLUSIONS AND DISCUSSION

We introduce a method of importance sampling applicable to spatial infection processes where observations of the numbers of infected individuals are made at intervals. This method provides estimates of parameters

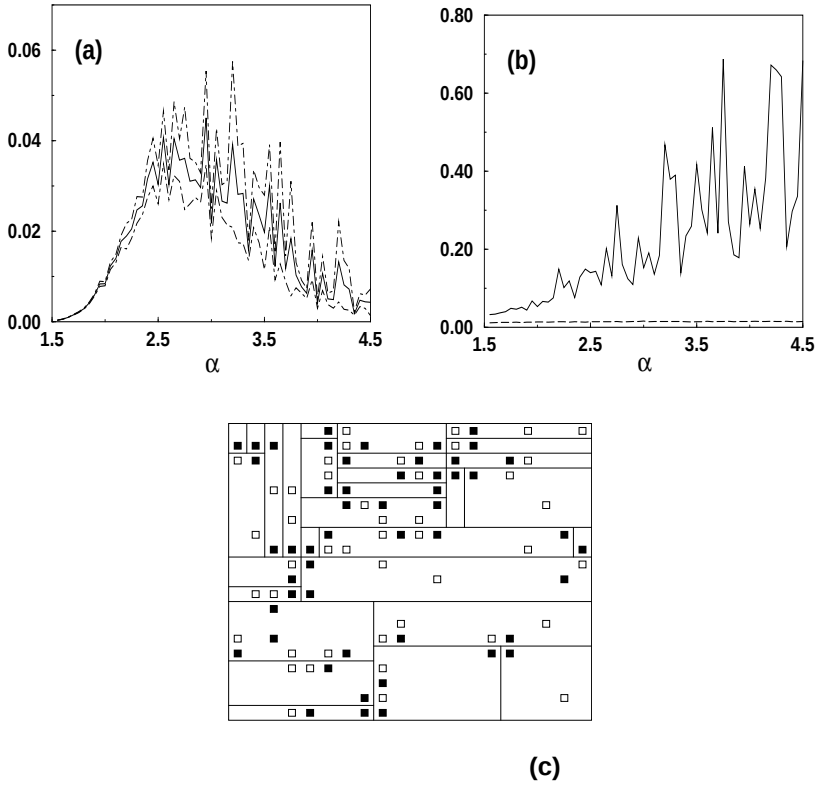

FIG. 3: Parameter likelihoods estimated from data generated by the infection process (1) with a power-law contact rate (3) with α = 2.5 (as used in Figure 2). (a) shows the estimate of L($\alpha \mid E$) (solid curve) and $\pm$ two standard deviation interval (dot-dashed curves) obtained by importance sampling using regions depicted in (c) containing no neighbouring secondary infections (see text for details), whilst (b) compares the coe cient of variance of the estimates of L($\alpha \mid E$) obtained (solid curve) with those of the importance sampling described in Figure 2 (dot-dashed curve). In (c) the open squares show the relative positions of the primary infections, the solid squares those of the secondary infections, and the lines delineate the regions used in the importance sampling algorithm

for the process (1) with contact distributions given by (3) for both model generated data, and for field observations. Parameter estimation based on the model data shows that, in accord with expectations, importance sampling is most effective in cases where the interactions are predominantly short-range. The utility of the method is further demonstrated through successful application to observations previously analysed using MCMC techniques by Gibson (1997a).

An important issue is the design of the regions used in the importance sampling procedure. Intuitively we should like borders to avoid going close to new infections if possible, and we have seen that regions must enclose sizeable fractions of any natural clusters in order to gain a worthwhile reduction in variance. However, if long chains of infection develop, and the features in the set of new infections are large (i.e. comparable to the number of nodes in the lattice) then it will not be possible to partition the lattice such that the assumption of independence of new infections between sectors is valid. The methodology developed here is therefore limited to time intervals for which the process has not has been able to establish such large structures. However, the application to the Citrus tristeza virus data set demonstrates that such time scales are of practical relevance (see also Gottwald et al. 1999). The concept of the identification of a critical scale for the process, as discussed by Rand and Wilson (1995), is relevant. This is a spatial scale defining regions over which development of the process is independent of that in other regions. It may be possible to link the variance reduction obtained with our importance sampling method to the problem of determining this critical scale, though this remains the subject of future work.

APPENDIX

This appendix shows how to calculate the importance sampling probability function $g(\omega)$ defined in Section 2. Consider the probability

$$\phi_j(\omega^j) = \Pr\{\text{the first } |X_j| \text{ infections (after } t_0) \text{ in } L_j \text{ occur at } X_j \text{ in the order } \omega^j \mid M'\}$$

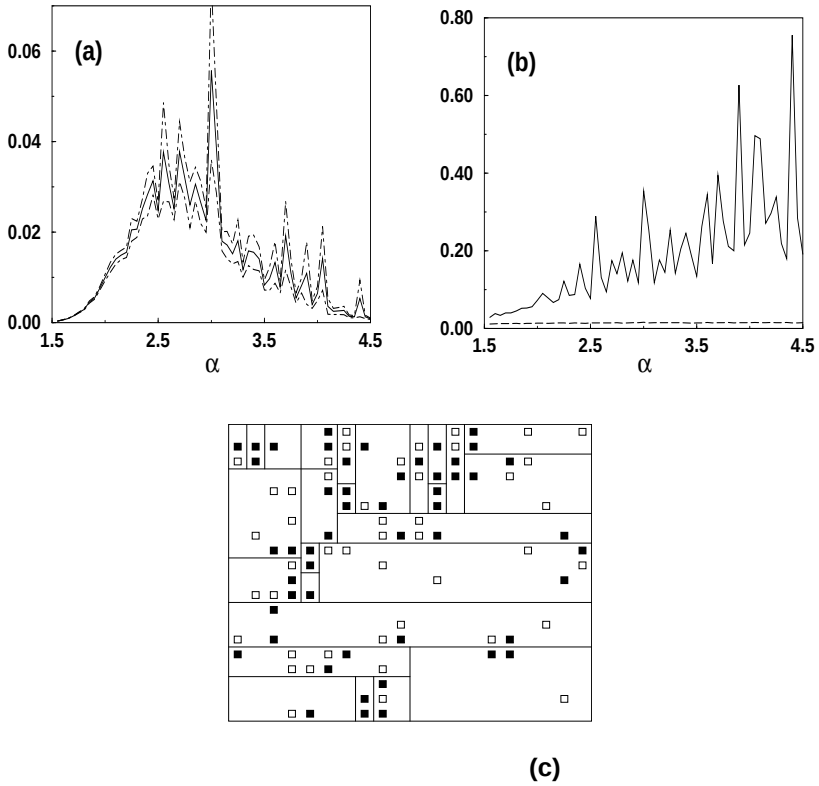

FIG. 4: As Figure 3, but with regions shown in (c) enclosing at most pair clusters of secondary infectives.

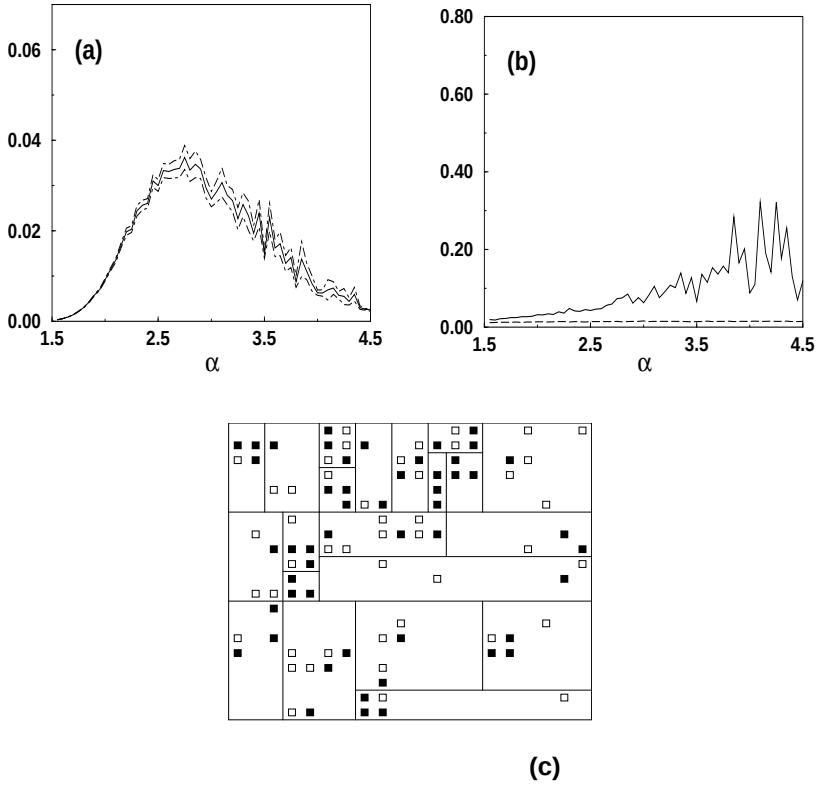

FIG. 5: As Figure 3, but with regions shown in (c) enclosing at most triple clusters of secondary infectives.

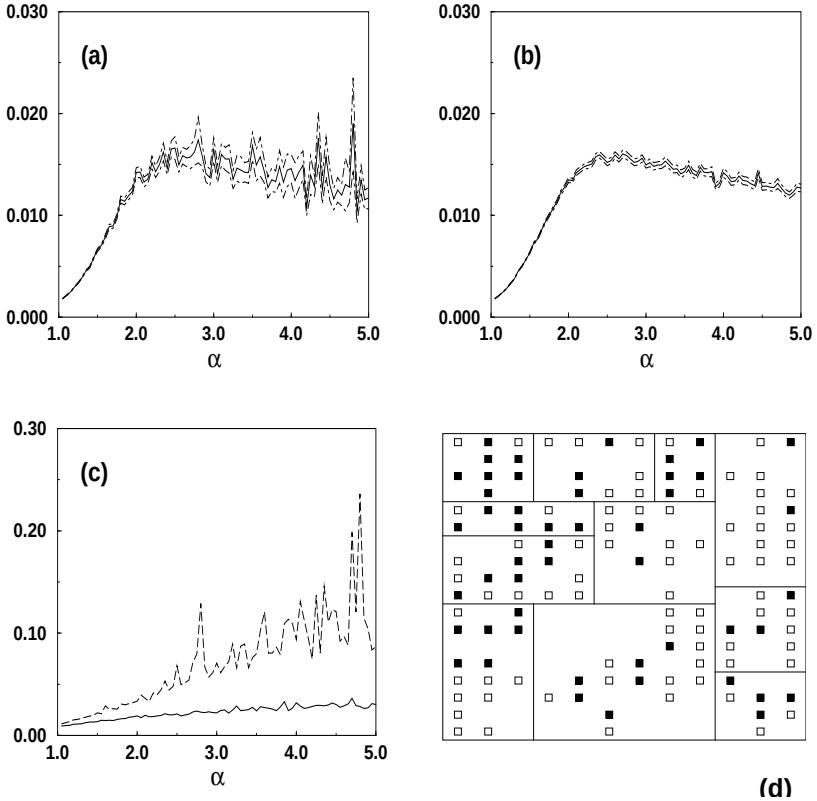

FIG. 6: Parameter likelihoods estimated from CTV data. (a) shows the estimate of $\mathsf{L}(\alpha \mid \beta = 0.1, E)$ (solid curve) and $\pm$ two standard deviation interval (dot-dashed curves) obtained by random sampling, whilst (b) shows that obtained by using the importance sampling procedure (see text for details). (c) shows the coe cient of variance of the estimates of $\mathsf{L}(\alpha \mid \beta = 0.1, E)$ obtained for random (solid curve) and importance (dot-dashed curve) sampling. In (d) the open squares show the relative positions of the primary infections, the solid squares those of the secondary infections, and the lines delineate the regions used in the importance sampling algorithm.

Note that for a single region $L_j = L$ covering the entire lattice this corresponds to the probability $\Pr(\omega \mid \alpha)$. In general, let n_j denote the number of secondary infections $|X_j|$ and $\{\boldsymbol{x_1}, ..., \boldsymbol{x_{n_j}}\}$ the locations of the ordering ω^j of the secondary infections X_j in subset L_j. Then

$$\phi_j(\omega^j) \;=\; \prod_{i=1}^{n_j} p_j(\boldsymbol{x_i} \mid \boldsymbol{x_1}, ..., \boldsymbol{x_{i-1}}, I_0).$$

Here the probability $p_j(\boldsymbol{x_i} \mid \boldsymbol{x_1}, ..., \boldsymbol{x_{i-1}}, I_0)$, that $\boldsymbol{x_i}$ is the i^{th} infection in the subset L_j, is dependent on the primary infections across the entire lattice I_0 but only on the $(i-1)$ secondary infections $\{\boldsymbol{x_1}, ..., \boldsymbol{x_{i-1}}\}$ in L_j. This may be expressed as the normalised infection pressure

$$p_j(\boldsymbol{x_i} \mid \boldsymbol{x_1}, ..., \boldsymbol{x_{i-1}}, I_0) \;= \frac{c(\boldsymbol{x_i} \mid \boldsymbol{x_1}, ..., \boldsymbol{x_{i-1}}, I_0)}{\sum_{\boldsymbol{x_k} \in L_j \setminus \{\boldsymbol{x_1}, ..., \boldsymbol{x_{i-1}}, I_0\}} c(\boldsymbol{x_k} \mid \boldsymbol{x_1}, ..., \boldsymbol{x_{i-1}}, I_0)},$$

where the summation is over all sites within L_j that remain susceptible following the infection at $\boldsymbol{x_{i-1}}$. The infection pressure at $\boldsymbol{x_i}$ is given by

$$c(\boldsymbol{x_i} \mid \boldsymbol{x_1}, ..., \boldsymbol{x_{i-1}}, I_0) = \sum_{\boldsymbol{y} \in \{\boldsymbol{x_1}, ..., \boldsymbol{x_{i-1}}, I_0\}} F_\alpha(\boldsymbol{x_i} - \boldsymbol{y}).$$

where $\boldsymbol{y}$ ranges over all primary infections in I_0 and the secondary infections $\boldsymbol{x_1}, ..., \boldsymbol{x_{i-1}}$. In order to generate a total ordering ω, sub-orderings ω^j are first generated for each subset L_j with probability $\phi_j(\omega^j)$. Denote by $n_{j,r}$ the number of unselected secondary infections in subset L_j, immediately prior to the selection of the r^{th} secondary infection in the total ordering ω. Thus, for the selection of the first secondary infection, $n_{j,r} = n_{j,1}$ corresponds to the number of secondary infectives n_j in L_j. If region j is selected in the r^{th} iteration then the secondary infective $(n_j - n_{j,r} + 1)$ in the sub-ordering ω^j is assigned to the r^{th} position in the total ordering ω, and $n_{j,r+1} = n_{j,r} - 1$. At each step the value of j is selected randomly with probability proportional to the number of unselected sites in L_j. In this way a sequence of subsets $\boldsymbol{J} = (j_1, ..., j_n)$, where $j_r \in \{1, 2, ..., k\}$, is generated with

probability

$$f(\boldsymbol{J}) = \prod_{r=1}^{n} \frac{n_{j,r}}{\sum_{j=1}^{k} n_{j,r}},$$

and the total ordering ω is generated with probability $g(\omega)$ as defined in Section 2.

REFERENCES

Cowles M.K. & Carlin B.P. 1996. Markov chain Monte Carlo convergence diagnostics: A comparative review, Journal of the American Statistical Society 91, 883-904.

Filipe J.A.N. & Gibson G.J. 1998. Studying and approximating spatio-temporal models for epidemic spread and control, Philiosophical Transactions Royal Society London B 353, 2153-2162.

Gibson G.J. 1997a. Investigating mechanisms of spatio-temporal epidemic spread using stochastic models, Phytopathology 87, 138-146.

Gibson G.J. 1997b. Markov Chain Monte Carlo methods for fitting and testing spatio-temporal stochastic models in plant epidemiology, Applied Statistics 46, 215-233.

Gibson G.J. and Austin E.J. 1996. Fitting and testing spatio-temporal stochastic models with application in plant epidemiology, Plant Pathology 45, 172-184.

Gibson G.J. and Renshaw E. 2001. Likelihood estimation for stochastic compartmental models using Markov chain methods, Statistics and Computing 11, 335-346.

Gottwald T. R., Gibson G. J., Garnsey S. M. and Irey M. 1999. Examination of the effect of aphid vector population composition on the spatial dynamics of citrus tristeza virus spread via stochastic modelling, Phytopathology 89, 603-608.

Marcus R., Svetlana F., Talpaz H., Salomon R. and Bar-Joseph M. 1982. On the spatial distribution of citrus tristeza virus disease, Phytoparasitica 12, 45-52.

O'Neill P.D. and Roberts G.O. 1999. Bayesian inference for partially observed stochastic epidemics. Journal of the Royal Statistical Society A 162, 121-129.

Propp J.G., and Wilson D.B. 1996. Exact sampling with coupled Markov chains and application to statistical mechanis. Random Structures Algorithms, 9, 223-252.

Rand D. A. and Wilson H. B. 1995. Using spatio-temporal chaos and intermediate-scale determinism to quantify spatially-extended ecosystems, Proceedings Royal Society. London. B 259, 111-117.

Ripley B. D. 1987. Stochastic Simulation. New York: Wiley.

Roberts C. P. and Cassella G. 1999. Monte Carlo Statistical Methods. New York: Springer.