Computational Statistics and Data Analysis 80 (2014) 240–250

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Efficient MCMC for temporal epidemics via parameter reduction

Fei Xiang [*], Peter Neal

Department of Mathematics and Statistics, Fylde College, Lancaster University, Lancaster, LA1 4YF, UK

ARTICLE INFO

Article history:
Received 13 January 2014
Received in revised form 19 May 2014
Accepted 6 July 2014
Available online 14 July 2014

Keywords:
SIR epidemic models
Data augmentation
Adaptive MCMC
Smallpox
Foot-and-mouth disease

ABSTRACT

An efficient, generic and simple to use Markov chain Monte Carlo (MCMC) algorithm for partially observed temporal epidemic models is introduced. The algorithm is designed to be adaptive so that it can easily be used by non-experts. There are two key features incorporated in the algorithm to develop an efficient algorithm, parameter reduction and efficient, multiple updates of the augmented infection times. The algorithm is successfully applied to two real life epidemic data sets, the Abakaliki smallpox data and the 2001 UK foot-and-mouth epidemic in Cumbria.

1. Introduction

Markov Chain Monte Carlo (MCMC) methods have been widely used for the statistical analysis of epidemic data, see, for example, Gibson (1997), O'Neill and Becker (2001), O'Neill (2009) and Britton et al. (2011). A major reason why MCMC has proved popular for epidemic models is that the data is necessarily, almost always incomplete. For example, with temporal epidemic data, we may know when individuals first show symptoms of a disease, but not when the individual was infected, or became infectious. Therefore to fit a parametric epidemic model with parameters, θ, to observed epidemic data $\mathbf{x}$, e.g., the removal time of infected individuals, large scale data augmentation is required. That is, the likelihood $\pi(\mathbf{x}|\theta) = \int \pi(\mathbf{x}, \mathbf{y}|\theta)\, dy$ is not in a tractable form for straightforward computation of the posterior distribution $\pi(\theta|\mathbf{x})$. However, given additional information (data), $\mathbf{y}$, e.g., the infection times of the individuals during the course of the epidemic, $\pi(\mathbf{x}, \mathbf{y}|\theta)$, is in a tractable form for analysis. Then the joint posterior distribution of $\mathbf{y}$ and θ given $\mathbf{x}$ satisfies

$$\pi(\theta, \mathbf{y}|\mathbf{x}) \propto \pi(\mathbf{x}, \mathbf{y}|\theta)\pi(\theta),$$

where $\pi(\theta)$ denotes the prior on θ. Throughout this paper the observed data, $\mathbf{x}$, are the removal times of infectious individuals and the augmented data, $\mathbf{y}$, are the infection times of individuals, and MCMC will be used to obtain samples from $\pi(\theta, \mathbf{y}|\mathbf{x})$.

The key MCMC question is how to efficiently sample from $\pi(\theta, \mathbf{y}|\mathbf{x})$, given that we are primarily interested in $\pi(\theta|\mathbf{x})$. In many circumstances the natural approach is to construct an MCMC algorithm that alternates between updating $\theta|\mathbf{y}, \mathbf{x}$ and updating $\mathbf{y}|\theta, \mathbf{x}$. This is sometimes referred to as a centred parameterisation, see Papaspiliopoulos et al. (2003). The centred parameterisation often performs poorly for epidemic models due to the strong dependence between the augmented data, $\mathbf{y}$

[*] Corresponding author. Tel.: +44 07942724271.
E-mail addresses: f.xiang@lancaster.ac.uk (F. Xiang), p.neal@lancaster.ac.uk (P. Neal).

http://dx.doi.org/10.1016/j.csda.2014.07.002
0167-9473/

and the model parameters θ, and therefore the use of (partially) non-centred algorithms have been advocated for epidemic models, see, for example, Neal and Roberts (2005), O'Neill (2009) and Jewell et al. (2009).

The aim of this paper is to develop an efficient, generic and simple to use MCMC algorithm that can be applied to partially observed temporal epidemic data sets. The algorithm is designed to be adaptive and to automatically tune itself so that the only inputs required from the user are the model specification and the data. This enables non-statisticians to utilise the algorithm. There are three key elements in obtaining an efficient MCMC algorithm. Firstly, parameter reduction by integrating out parameters following Liu (1994). That is, we write $\theta = (\phi, \omega)$ and compute $\pi(\phi, \mathbf{y}|\mathbf{x}) = \int \pi(\phi, \omega, \mathbf{y}|\mathbf{x}) d\omega$. We then construct an MCMC algorithm that obtains samples from $\pi(\phi, \mathbf{y}|\mathbf{x})$. Given a sample $(\phi_1, \mathbf{y}_1), (\phi_2, \mathbf{y}_2), \ldots, (\phi_n, \mathbf{y}_n)$ from $\pi(\phi, \mathbf{y}|\mathbf{x})$, a sample from $\theta_1, \theta_2, \ldots, \theta_n$ can be obtained from $\pi(\theta|\mathbf{x})$ by sampling ω_i from $\pi(\omega_i|\phi_i, \mathbf{y}_i, \mathbf{x})$. This approach has previously been used for epidemic models in Neal and Roberts (2005) and Jewell et al. (2009). Secondly, we introduce a novel data driven updating scheme which is applicable when the parameters, ω, integrated out of the model are ideally wanted to propose new values for $\mathbf{y}$. That is, we look to incorporate the efficiency gains of integrating out parameters with the advantages of proposing $\mathbf{y}|\theta, \mathbf{x}$ in the MCMC algorithm. Thirdly, we introduce an adaptive scheme for choosing the number of components of $\mathbf{y}$ to update in each iteration. This is an important contribution in ensuring that the resulting MCMC algorithm effectively explores $\pi(\theta, \mathbf{y}|\mathbf{x})$. The key contribution of the paper is to combine the above features to give a simple to use, adaptive and data driven MCMC algorithm which is applicable for a wide range of SIR (Susceptible → Infective → Removed) epidemic models, whose performance is at least comparable with state-of-the-art MCMC algorithms for partially observed temporal epidemic data.

The paper is structured as follows. In Section 2, we give a generic description of the (SIR) epidemic model and present a detailed motivation for the approach taken in this paper. In Section 3, we develop the MCMC methodology for the homogeneously mixing epidemic to illustrate its applicability and to allow for comparisons with other approaches, in particular, the partial non-centred MCMC algorithm of Neal and Roberts (2005). In Section 4, we apply the generic algorithm to a spatial epidemic outbreak, the 2001 UK foot-and-mouth disease (FMD) outbreak in Cumbria. This is a substantial outbreak with 1021 farms out of 5378 farms infected with FMD. The algorithm is shown to perform well in this case and this highlights the general applicability of the methodology. Finally in Section 5, we briefly discuss extensions of the current work.

2. SIR epidemic model

In this section we give the generic framework that we shall be considering in this paper. We assume that there is one introductory infectious case which introduces the disease into the population. (The extension to multiple introductory cases is trivial.) The population is assumed to be a closed population of N individuals, labelled $1, 2, \ldots, N$, and there is assumed to be no births, deaths, immigration or emigration during the course of the epidemic. This is a reasonable assumption for an epidemic outbreak in a small community, in that, the time course of the epidemic is likely to be in the order of weeks and during this period it is highly unlikely that the population will experience major demographic changes. We assume that the disease follows an *SIR* epidemic model, where initially all individuals, except the introductory case, are susceptible. On becoming infectious, an individual is infectious for a given period of time, distributed according to a random variable Q. We take $Q \sim \text{Gamma}(\alpha, \delta)$ with probability density function $f_Q(x) = \delta^\alpha x^{\alpha-1} \exp(-\delta x)/\Gamma(\alpha)$, although as demonstrated in the Supplementary material (see Appendix A) with Q following a Weibull distribution, the methods are appropriate for other families of parametric distribution. We assume that whilst infectious an individual i makes infectious contacts at the points of a homogeneous Poisson point process with rate $\sum_k \lambda_{ik}$. The probability that an infectious contact made by individual i is with individual j is $\lambda_{ij}/\sum_k \lambda_{ik}$. Therefore the probability that, during its infectious period of length Q_i, individual i makes at least one infectious contact with individual j is $1 - \exp(-\lambda_{ij} Q_i)$. At this stage we assume a general form for the contact rate between individuals i and j and, for each example we will assume a parametric form for $\lambda_{ij}(\vartheta)$ depending upon properties of individuals i and j, the relationship (*e.g.* spatial distance) between individuals i and j and the parameters ϑ. An infectious contact made by an infective with a susceptible results in infection of the individual, whereas an infectious contact with a non-susceptible has no effect on the recipient. At the end of the infectious period, the individual recovers from the disease and is immune to further infection. It is straightforward to extend the model to an *SEIR* epidemic model which incorporates an exposed state, where individuals are infected but not yet infectious.

Throughout we shall consider completed epidemics, that is, we assume that m individuals have been infected during the course of the epidemic which has now finished. However there is nothing in the methodology which limits it to this case and it could easily be applied to epidemics in progress. For simplicity, we label the m infected individuals $i = 1, 2, \ldots, m$ and the $N - m$ individuals who remain susceptible, we label $i = m+1, m+2, \ldots, N$. For an individual, i say, infected during the course of the epidemic, we need to know the point of time at which they became infected, I_i, and the time at which they recovered/became removed, R_i. Typically, it is assumed that R_i is observed and this will often correspond to the appearance of symptoms of an individual. However, the time at which an individual becomes infected is rarely known and therefore we assume that I_i is unknown.

We are now in position to construct the likelihood. For convenience we assume that the infected individuals are labelled according to their removal times such that $R_1 \leq R_2 \leq \cdots R_m$. Let $\mathbf{R} = (R_1, R_2, \ldots, R_m)$ and $\mathbf{I} = (I_1, I_2, \ldots, I_m)$ denote the set of removal and infection times, respectively. Let θ denote the parameters of the model with θ comprising the infectious period parameters (α, δ) and the infectious process parameters ϑ. Let κ denote the index of the initial infective such that $I_\kappa = \min\{I_j; j = 1, 2, \dots, m\}$. The likelihood of the data satisfies

$$\pi(\mathbf{R}, \mathbf{I}|\theta) = \prod_{j \neq \kappa} \left\{ \sum_{k: I_k < I_j \leq R_k} \lambda_{kj}(\vartheta) \right\} \exp\left(- \sum_{k=1}^{m} \sum_{j=1}^{N} \lambda_{kj}(\vartheta) \{(R_k \wedge I_j) - (I_k \wedge I_j)\} \right) \times \prod_{i=1}^{m} \left\{ \frac{\alpha^\delta}{\Gamma(\alpha)} (R_i - I_i)^{\alpha - 1} \exp(-\delta\{R_i - I_i\}) \right\}, \quad (1)$$

where for notational convenience, we set $I_{m+1} = I_{m+2} = \cdots = I_N = \infty$ for those individuals who remain susceptible throughout the course of the epidemic. The above set up is very general and can be applied to a wide range of population models. For example, for the homogeneously mixing epidemic, $\lambda_{kj}(\vartheta) = \beta$ (Section 3) and for a spatial epidemic, $\lambda_{kj}(\vartheta) = \beta \exp(-\psi d(k, j))$, where $d(k, j)$ denotes the distance between individuals k and j (an extension of this model is used in Section 4). Other prime examples in the literature are the household epidemic model where $\lambda_{kj}(\vartheta) = \beta_H$, if individuals k and j belong to the same household and $\lambda_{kj}(\vartheta) = \beta_G$ otherwise (O'Neill, 2009) and the network epidemic model where $\lambda_{kj}(\vartheta) = \beta_N$ if an edge exists between individuals k and j in the underlying network and $\lambda_{kj}(\vartheta) = 0$ otherwise (Britton and O'Neill, 2002).

Let $\pi(\theta) = \pi(\vartheta)\pi(\alpha)\pi(\delta)$ denote the prior on the parameters. For α and δ we assume a Gamma(a_ν, b_ν) prior, where $\nu = \alpha, \delta$, respectively and the prior on ϑ will be problem specific. Then the joint distribution of the parameters θ and the unobserved infection times $\mathbf{I}$ given the observed removal times $\mathbf{R}$ satisfies

$$\pi(\mathbf{I}, \theta|\mathbf{R}) \propto \prod_{j \neq \kappa} \left\{ \sum_{k: I_k < I_j \leq R_k} \lambda_{kj}(\vartheta) \right\} \exp\left(- \sum_{k=1}^{m} \sum_{j=1}^{N} \lambda_{kj}(\vartheta) \{(R_k \wedge I_j) - (I_k \wedge I_j)\} \right) \times \prod_{i=1}^{m} \left\{ \frac{\delta^\alpha}{\Gamma(\alpha)} (R_i - I_i)^{\alpha - 1} \exp(-\delta\{R_i - I_i\}) \right\} \times \pi(\vartheta)\alpha^{a_\alpha - 1} \exp(-\alpha b_\alpha)\delta^{a_\delta - 1} \exp(-\delta b_\delta). \quad (2)$$

It is then straightforward to set up an MCMC algorithm to obtain samples from $\pi(\theta|\mathbf{R})$, the posterior distribution of the parameters given the observed removal times using (2). This can be done by using a data augmentation, Metropolis-within-Gibbs algorithm which alternates between updating $\mathbf{I}$ given θ and $\mathbf{R}$ and updating θ given $\mathbf{I}$ and $\mathbf{R}$. The key question is, how to do this in an efficient manner?

The simplest MCMC scheme is to alternate between updating the infection times $\mathbf{I}$ and parameters θ. Typically, updating of the infection times is done one at a time, either a randomly chosen infection time Neal and Roberts (2005) or all infection times sequentially Jewell et al. (2009) per MCMC iteration. Then the components of θ are updated individually in a sequential order. The conditional distribution of I_i given the remainder of the data and the parameters is not of a convenient form for a Gibbs step. However, we know that the infectious period distributions follow Gamma(α, δ) and therefore it is convenient to use an independent sampler for I_i, proposing $I_i' = R_i - \text{Gamma}(\alpha, \delta)$, see, for example, Neal and Roberts (2005). For the components of θ it is possible in some cases to use a Gibbs step, since for example, from (2),

$$\delta|\{\mathbf{R}, \mathbf{I}, \vartheta, \alpha\} \sim \text{Gamma}\left(m\alpha + a_\delta, \sum_{i=1}^{m} (R_i - I_i) + b_\delta \right). \quad (3)$$

However, for example, α does not have a nice standard conditional distribution and it is necessary to use an alternative such as random walk Metropolis (RWM).

The above MCMC algorithm is easy to implement but not very efficient and if m is large and all infection times are updated per iteration, it can take a long time per iteration. The key problem is the intrinsic dependence between $\mathbf{I}$ and θ, see Neal and Roberts (2005). A solution proposed by Neal and Roberts (2005) and subsequently used in Jewell et al. (2009), is a non-centred parameterisation. That is, to reparameterise the model to express the infectious period of individual i as $R_i - I_i = U_i/\delta$, or equivalently to write $I_i = R_i - U_i/\delta$, where $U_i \sim \text{Gamma}(\alpha, 1)$ and U_i is *a priori* independent of δ. Then keeping $\mathbf{U} = (U_1, U_2, \dots, U_m)$ fixed but proposing a new δ (possibly using a random walk Metropolis proposal) updates all the infection times $\mathbf{I}$. However such a process maintains a given ratio between infectious periods and a partial non-centred algorithm, where only some of the infection times change with δ is found to work best. The partial non-centred MCMC algorithm is found to work well in Neal and Roberts (2005) and Jewell et al. (2009) but it is non-trivial to optimise the algorithm in terms of both tuning the proposal distribution for updating δ and the proportion of infection periods to non-centre.

In this paper we look to introduce an efficient MCMC updating schema for epidemic models by placing particular emphasis on the updating of $\mathbf{I}$. There are a myriad of MCMC algorithms which could be used to update $\mathbf{I}$ but a common effective approach is to propose $I_i \sim R_i - \text{Gamma}(\alpha, \delta)$ (Neal and Roberts, 2005; Jewell et al., 2009). In particular, we look at how this proposal scheme can be implemented with δ integrated out and how many components of $\mathbf{I}$ to propose to update at a given time. From (2), we observe that ϑ and (α, δ) are conditionally independent given $\mathbf{I}$ and that the different parameters depend upon **I** in different ways. Thus if we consider the marginal posterior distributions of different parameters it is likely that different MCMC schemes are going to be optimal in obtaining samples from the marginal posterior distributions of different parameters. Consequently, our aim is for a mechanism for automatically tuning the MCMC algorithm in such a way that the algorithm performs well for all the parameters in the model. To this end we focus on $B = \sum_{i=1}^{m}(R_i - I_i)$, the sum of the infectious periods, as the key summary statistic for **I** and seek to optimise the mixing of the MCMC algorithm with respect to B. Note that the conditional distribution of δ given by (3) is intrinsically linked to B, and optimising the mixing in B optimises the mixing in δ. The key ideas behind efficient exploration of the space **I** are as follows. Firstly, the focus is on the augmented data **I** rather than the parameters θ. Therefore we follow Neal and Roberts (2005) in seeking to integrate out parameters (Liu, 1994) to improve the performance of the MCMC algorithm and allow more freedom in the updating of **I**. Estimates of summary statistics of the posterior distribution of the integrated out parameters can be obtained in a straightforward manner. Secondly, we explore the question of how many components of **I** to update at each iteration. This is similar to the question of the optimal partial non-centring in Neal and Roberts (2005). A key difference to Neal and Roberts (2005) is that we introduce a mechanism for the data to automatically tune the number of components of **I** updated. Moreover, in Section 3, we discuss the benefits of not updating the same number of components at each iteration but instead drawing the number of components to be updated from a probability distribution which is determined by the data.

In introducing the above innovations it is important to be able to compare the performance of updating different number of components of **I** and more generally the performance of different MCMC algorithms. A key tool to compare MCMC algorithms is to measure the serial autocorrelation in the estimation of B and to compute the effective sample size for B of the sample obtained from the MCMC algorithm. A useful proxy for these are the lag-1 autocorrelation which is usually an excellent guide of MCMC performance and for the random walk Metropolis (RWM) it is known that for a large number of parameters, k, (specifically in the limit as $k \to \infty$) minimising the lag-1 autocorrelation optimises the performance of the RWM algorithm (see, Roberts et al., 1997, Roberts and Rosenthal, 1998). In turn the lag-1 autocorrelation is minimised by maximising the expected squared jumping distance $\mathbb{E}[(B_1 - B_0)^2]$, where B_j denotes the sum of the infectious periods at the end of the j th iteration and B_0 is drawn from the stationary distribution of B. For assessing MCMC performance for updating different numbers of infection times, we use a simple measure which has high positive correlation with $\mathbb{E}[(B_1 - B_0)^2]$. For assessing the performance of proposing to update p infection times, we compute $p\text{Acc}_p$, where Acc_p is the mean proportion of proposed moves accepted with p infection times updated.

We proceed by implementing these ideas with two examples. In Section 3, we consider a simple homogeneously mixing epidemic model. This example allows us to study the behaviour and performance of the algorithm in detail. Then in Section 4 we apply the methodology to a subset of the 2001 UK Foot-and-Mouth disease (FMD) epidemic outbreak, specifically the outbreak in Cumbria which saw 1021 farms infected with FMD. This was one of the most severely hit regions of the UK.

3. Homogeneously mixing epidemic

The simplest model satisfying (1) is the homogeneously mixing epidemic model where for all $k, j = 1, 2, \ldots, N$, $\lambda_{kj}(\vartheta) = \beta$. In that case, letting $B = \sum_{i=1}^{m}(R_i - I_i)$ and

$$\begin{aligned} A &= \sum_{k=1}^{m}\sum_{j=1}^{N}\{(R_k \wedge I_j) - (I_k \wedge I_j)\} \\ &= \sum_{k=1}^{m}\sum_{j=1}^{m}\{(R_k \wedge I_j) - (I_k \wedge I_j)\} + (N-m)B, \end{aligned} \tag{4}$$

we can rewrite (1) as

$$\pi(\mathbf{R}, \mathbf{I}|\alpha, \delta, \beta) = \prod_{j \neq \kappa}\left\{\beta Y_{I_j -}\right\}\exp(-\beta A) \times \prod_{i=1}^{m}\left\{\frac{\alpha^{\delta}}{\Gamma(\alpha)}(R_i - I_i)^{\alpha-1}\right\}\exp(-\delta B), \tag{5}$$

where Y_t denotes the total number of infectives at time t and $Y_{t-} = \lim_{s\uparrow t} Y_s$, the total number of infectives just prior to time t.

Suppose that we assume independent gamma priors for each of the parameters with Gamma(a_ν, b_ν) denoting the prior for parameter ν and $\nu = \alpha, \delta, \beta$. Unless otherwise stated we assume Gamma(1, 1) prior distribution for all parameters. Then using (4) and following (2), it is straightforward to show that the joint distribution of θ and **I** given **R** satisfies

$$\begin{aligned} \pi(\mathbf{I}, \theta|\mathbf{R}) \propto\ & \beta^{m-1}\prod_{j \neq \kappa} Y_{I_j -}\exp(-\beta A) \times \prod_{i=1}^{m}\left\{\frac{\alpha^{\delta}}{\Gamma(\alpha)}(R_i - I_i)^{\alpha-1}\right\}\exp(-\delta B) \\ & \times \beta^{a_\beta - 1}\exp(-\beta b_\beta)\alpha^{a_\alpha - 1}\exp(-\alpha b_\alpha)\delta^{a_\delta - 1}\exp(-\delta b_\delta). \end{aligned} \tag{6}$$

The key step as highlighted in Section 2 is to construct an MCMC algorithm that effectively explores the space of infection times **I**. To this end, we follow Neal and Roberts (2005), in integrating out parameters, where possible, to construct an MCMC algorithm on a smaller target space. It is trivial to integrate out β and δ to give

$$f(\mathbf{I}, \alpha|\mathbf{R}) \propto \prod_{j\neq\kappa}\{Y_{I_j-}\}\frac{\Gamma(m+a_\beta-1)}{(b_\beta+A)^{m+a_\beta-1}} \times \prod_{i=1}^{m}(R_i-I_i)^{\alpha-1}\frac{\Gamma(m\alpha+a_\delta)}{(b_\delta+B)^{m\alpha+a_\delta}} \times \frac{\alpha^{a_\alpha-1}\exp(-b_\alpha\alpha)}{\Gamma(\alpha)^m}. \tag{7}$$

Note that in Neal and Roberts (2005) only β is integrated out as δ is needed for the non-centred algorithm. Integrating out δ, is at first glance problematic, in that, it inhibits the use of $R_i - \text{Gamma}(\alpha, \delta)$ as a proposal for I_i. We give details below on how this is circumvented.

We proceed by detailing an MCMC algorithm for obtaining samples from $(\mathbf{I}, \alpha)$. It should be noted that samples from β and δ can easily be obtained since $\pi(\beta|\mathbf{I}, \mathbf{R}, \alpha) \sim \text{Gamma}(m+a_\beta-1, b_\beta+A)$ and $\pi(\delta|\mathbf{I}, \mathbf{R}, \alpha) \sim \text{Gamma}(m\alpha+a_\delta, b_\delta+B)$, exploiting the conditional independence of β and δ given $(\mathbf{I}, \alpha)$. The MCMC algorithm we implement alternates between updating $\mathbf{I}|\alpha, \mathbf{R}$ and $\alpha|\mathbf{I}, \mathbf{R}$ when α is unknown and simply consists of $\mathbf{I}|\alpha, \mathbf{R}$ updates in the case α is known. An example of a case where α can be assumed to be known is the general stochastic epidemic model, see Bailey (1975) and O'Neill and Roberts (1999), where $\alpha = 1$ and the infectious periods are exponentially distributed. In the case of known α we simply replace the gamma distributed prior by a point mass distribution at the given value of α.

We proceeding by describing the updating schema for $\mathbf{I}|\alpha, \mathbf{R}$.

Updating I

1. Choose to update p infectious periods with probability u_p ($\sum_{j=1}^m u_j = 1$). (We discuss the choice of $\mathbf{u} = (u_1, u_2, \dots, u_m)$ below.) Choose a set $\mathbf{k} = \{k_1, k_2, \dots, k_p\}$ uniformly, at random, from $\{1, 2, \dots, m\}$. The probability of selecting $\mathbf{k}$, given p, is $1/\binom{m}{p}$.
2. Draw γ from $\text{Gamma}(\alpha m + a_\delta, b_\delta + B)$.
   That is, we proposed γ from the conditional posterior distribution of δ.
3. Draw $Q'_{k_1}, Q'_{k_2}, \dots, Q'_{k_p}$ independently from $\text{Gamma}(\alpha, \gamma)$ and for $j = 1, 2, \dots, p$, set $I'_{k_j} = R_{k_j} - Q'_{k_j}$.
   Since γ is drawn from the conditional posterior distribution of δ, we are essentially implementing the independent sampler used in, for example, Neal and Roberts (2005). The key difference is that γ is only introduced as an intermediary to facilitate the proposal of $\mathbf{Q}'$ and is integrated out in computing the proposal density.
4. Compute the proposal density, $h(\mathbf{I} \to \mathbf{I}')$ for proposing the move from $\mathbf{I}$ to $\mathbf{I}'$, or equivalently, from $\mathbf{Q}$ to $\mathbf{Q}'$,

$$h(\mathbf{Q} \to \mathbf{Q}') = u_p\frac{(m-p)!p!}{m!} \times \int_0^\infty f_{\mathbf{Q}'}(\mathbf{q}|\gamma)f_\gamma(\gamma|\mathbf{Q})\,d\gamma,$$

   where $q_{k_i}(=Q'_{k_i})$ is the proposed value for Q_{k_i} and

$$\begin{aligned}\int_0^\infty f_{\mathbf{Q}'}(\mathbf{q}|\gamma)f_\gamma(\gamma|\mathbf{Q})\,d\gamma &= \int_0^\infty \frac{(b_\delta+B)^{m\alpha+a_\delta}}{\Gamma(m\alpha+a_\delta)}\gamma^{m\alpha+a_\delta-1}\exp(-(b_\delta+B)\delta)\prod_{i=1}^{p}\left(\frac{\gamma^\alpha}{\Gamma(\alpha)}(Q'_{k_i})^{\alpha-1}\exp(-\gamma Q'_{k_i})\right)d\gamma \\ &= \frac{(b_\delta+B)^{m\alpha+a_\delta}}{\Gamma(\alpha)^p\Gamma(m\alpha+a_\delta)}\prod_{i=1}^{p}(Q'_{k_i})^{\alpha-1}\int_0^\infty \gamma^{p\alpha+n_I\alpha+a_\delta-1}\exp\left(-\gamma\left(B+b_\delta+\sum_{i=1}^{p}Q'_{k_i}\right)\right)d\gamma \\ &= \frac{(b_\delta+B)^{m\alpha+a_\delta}}{\Gamma(\alpha)^p\Gamma(m\alpha+a_\delta)}\prod_{i=1}^{p}(Q'_{k_i})^{\alpha-1}\frac{\Gamma(p\alpha+m\alpha+a_\delta)}{\left(B+b_\delta+\sum_{i=1}^{p}Q'_{k_i}\right)^{p\alpha+m\alpha+a_\delta}}.\end{aligned}$$

   Therefore

$$h(\mathbf{Q} \to \mathbf{Q}') = u_p\frac{(m-p)!p!}{m!} \times \frac{\Gamma(\alpha p+\alpha m+a_\delta)}{\Gamma(\alpha)^p\Gamma(\alpha m+a_\delta)} \times \frac{\prod_{i=1}^{p}(Q'_{k_i})^{\alpha-1}(B+b_\delta)^{m\alpha+a_\delta}}{\left(B+b_\delta+\sum_{i=1}^{p}Q'_{k_i}\right)^{(m+p)\alpha+a_\delta}}. \tag{8}$$

5. Compute the acceptance probability, Acc, for the proposed move from $\mathbf{I}$ to $\mathbf{I}'$, where

$$\text{Acc} = \min\left\{1, \frac{f(\mathbf{I}', \alpha|\mathbf{R})}{f(\mathbf{I}, \alpha|\mathbf{R})} \times \frac{h(\mathbf{Q}' \to \mathbf{Q})}{h(\mathbf{Q} \to \mathbf{Q}')}\right\}.$$

   It follows from (7) and (8) that, since α is fixed,

$$\text{Acc} = \frac{\prod_{j\neq\kappa'} Y_{I'_j-}}{\prod_{j\neq\kappa} Y_{I_j-}}\left(\frac{b_\beta+A}{b_\beta+A'}\right)^{m+a_\beta-1}\left(\frac{B'+b_\delta+\sum_{i=1}^{p}Q_{k_i}}{B+b_\delta+\sum_{i=1}^{p}Q'_{k_i}}\right)^{(m+p)\alpha+a_\delta}. \tag{9}$$

However, $B = \sum_{i=1}^{m} Q_i$ and it is trivial to show that the last term on the right hand side of (9) is 1, giving

$$\text{Acc} = \frac{\prod_{j \neq \kappa'} Y_{I'_j -}}{\prod_{j \neq \kappa} Y_{I_j -}} \left(\frac{b_\beta + A}{b_\beta + A'} \right)^{m + a_\beta - 1}. \tag{10}$$

We update α using random walk Metropolis algorithm (RWM) on the log scale. That is, we propose $\log \alpha' \sim N(\log \alpha, \sigma_\alpha^2)$. This was found to be more effective than using standard RWM. From (7) the conditional distribution of α given $\mathbf{I}$ is,

$$\pi(\alpha | \mathbf{I}, \mathbf{R}) \propto \prod_{i=1}^{m} (R_i - I_i)^{\alpha - 1} \frac{\Gamma(m\alpha + a_\delta)}{(b_\delta + B)^{m\alpha + a_\delta}} \times \frac{\alpha^{a_\alpha - 1} \exp(-b_\alpha \alpha)}{\Gamma(\alpha)^m}. \tag{11}$$

Hence, we accept α' with probability

$$\min\left(1, \frac{\alpha' \pi(\alpha' | \mathbf{I}, \mathbf{R})}{\alpha \pi(\alpha | \mathbf{I}, \mathbf{R})}\right).$$

The key question that needs to be addressed with the updating schema for $\mathbf{I}$ is, what is a good choice of distribution for $\mathbf{u}$? For example, should $u_p = 1$ for some $1 \leq p \leq m$, that is, there is an optimal number, p, of infectious periods to update per iteration or are there advantages to allowing p to vary? Also does the choice of $\mathbf{u}$ depend upon α. Finally, how efficient is the algorithm in updating α? To answer these questions, we apply the above MCMC algorithm to the Abakaliki smallpox data, Bailey (1975, p. 125). The population is a closed community of $N = 120$ individuals, of whom $m = 30$ become infected with smallpox. The data consists of 29 inter-removal times, measured in days:

13, 7, 2, 3, 0, 0, 1, 4, 5, 3, 2, 0, 2, 0, 5, 3, 1, 4, 0, 1, 1, 1, 2, 0, 1, 5, 0, 5, 5.

Thus $R_1 = 0, R_2 = 13, R_3 = 20, \ldots, R_{30} = 76 = T$. This data set has been studied extensively, see, for example, O'Neill and Roberts (1999), O'Neill and Becker (2001), Neal and Roberts (2005), McKinley et al. (2014), and is therefore useful for benchmarking of our approach.

The first step is to choose initial values for $\mathbf{I}$ and α. The key requirement is that the initial choice of $\mathbf{I}$ is consistent with the data, that is, there is always at least one infectious individual from the start (time I_κ) to the end (time $T = R_m$) of the epidemic. A straightforward approach is to simulate $I_i = R_i - \text{Gamma}(\alpha, \delta)(i = 1, 2, \ldots, m)$ for some predefined (α, δ) and if the resulting $\mathbf{I}$ is consistent with the data initiate the MCMC algorithm, otherwise re-simulate with possibly different (α, δ). For the Abakaliki data set, taking $\alpha = 1$ and $\delta = 0.1$ is consistent with previous analysis in O'Neill and Roberts (1999) and Neal and Roberts (2005) which fit the general stochastic epidemic model to the data. In general, the initial choice for (α, δ) is problem specific and will depend upon prior information about the disease, but setting $\alpha = 1$ and choosing a δ value such that both $\mathbf{I}$ is consistent with the data but only a relatively small proportion of infection times are prior to R_1 will work well. Note that choosing δ very small is likely to result in infectious periods $\mathbf{I}$ which are compatible with the data but highly unlikely in that most of the infections occur before time R_1.

We now address the question of how to choose $\mathbf{u}$, starting with the case α is fixed. Throughout we discuss our findings in relation to the Abakaliki data set but in all cases these are supported by extensive simulation studies. A natural starting point is to set $\mathbf{u} = \mathbf{v}_p$ $(1 \leq p \leq m)$, where the p^{th} component of $\mathbf{v}_p$ is 1 and all other components are equal to 0. For the Abakaliki data we considered all combinations of $p = 1, 2, \ldots, m$ in conjunction with $\alpha = 1, 4, 10$, running the MCMC algorithm for 3000 iterations and repeating 10 times. We observed that the acceptance probability is decreasing in p but generally increasing in α. For $p \geq 25$ and $\alpha \geq 4$, we observed the acceptance rate drops off significantly. This is due to the algorithm becoming stuck at the initial configuration of $\mathbf{I}$ for a large number of iterations. However, once the algorithm moves away from its initial starting point good mixing is observed for large values of p. This suggests that choosing a $\mathbf{u}$ which allows for different infection times to be updated could be best, in overcoming the possibly poor choice of starting values $\mathbf{I}$ and to allow for good mixing of $\mathbf{I}$ in stationarity. The optimal mean number of infection times updated was 2.5 with $p = 9$ for $\alpha = 1$, 5 with $p = 18$ for $\alpha = 4$ and 7 with $p = 19$ for $\alpha = 10$. This supports using a larger value of p as α increases and is to be expected as the infectious periods have smaller variance as α increases.

For fixed α, our findings suggests not fixing the number of infection times, p, to update and this is even more pertinent when α is unknown. Therefore we outline a generic approach for automatically tuning $\mathbf{u}$ and σ_α to obtain an efficient algorithm. We initialise $\mathbf{u} = (1/m, 1/m, \ldots, 1/m)$, an initial value for $\sigma_\alpha = 1$ and a fixed number of tuning iterations $S = 10\,000$. For the first S iterations, draw $p = j$ according to $\mathbf{u}$ and implement the updating scheme for $\mathbf{I}$ with $p = j$, recording whether or not the proposed move is accepted. After S iterations, for each $j = 1, 2, \ldots, m$, set $\tau_j = j\text{Acc}_j$, where Acc_j is the proportion of accepted moves with $p = j$. Then set $u_j \propto \tau_j^d$, for some $d > 0$. The larger d is the more weight is given to the large τ_j values.

For the RWM update of α it is known that it is close to optimal to choose σ_α such that approximately a quarter of proposed moves are accepted, see, for example, Roberts et al. (1997). Therefore we use the following adaptive procedure during the first S iterations to update σ_α, based on that used in Xifara (2013, Chapter 1). If a proposed move is accepted at iteration $J = 1, 2, \ldots, S$, we increase σ_α by setting

$$\sigma_\alpha^2 = \sigma_\alpha^2 + 3\frac{\sigma_\alpha^2}{100\sqrt{J}} \tag{12}$$

and if a proposed move is rejected at iteration $J = 1, 2, \ldots, S$, we decrease σ_α by setting

$$\sigma_\alpha^2 = \sigma_\alpha^2 - \frac{\sigma_\alpha^2}{100\sqrt{J}}. \tag{13}$$

Thus after S iterations, we have an improved $\mathbf{u}$ and σ_α. The above process can be repeated for a further S iterations, if further refinement of $\mathbf{u}$ and σ_α is required. Further refinement is likely if the initial value of σ_α is far too large or small.

In implementing the above adaptive algorithm on the Abakaliki small pox data set, we considered two scenarios. Firstly, $\alpha = 1$, which corresponds to the general stochastic epidemic model and allows for comparison with previous analysis in Neal and Roberts (2005). Secondly, α unknown with an Exp(0.001) prior on α. (The prior mean on α is 1000.) We ran the adaptive step twice with $d = 3$ and $S = 10\,000$ and we discarded these first 20 000 iterations as burn-in. The algorithm was then run for a further 100 000 iterations with the selected $\mathbf{u}$, and where appropriate, σ_α to obtain samples from $\pi(\alpha, \mathbf{I}|\mathbf{R})$, and hence, to obtain samples $\pi(\alpha, \beta, \delta|\mathbf{R})$. For unknown α, we obtained $\sigma_\alpha = 1.45$ with an acceptance rate of 27.47% which is close to optimal. For $\alpha = 1$, we found that $\sum_{i=6}^{15} u_i = 0.598$, that is, the majority of proposed moves involve updating between 6 and 15 infection times but that all $u_i > 0.001$ $(1 \leq i \leq 30)$. For comparing the performance of our algorithm with the non-centred algorithm on Neal and Roberts (2005), we generated $(\beta_1, \delta_1), \ldots, (\beta_{100\,000}, \delta_{100\,000})$ by sampling β_i and δ_i from the appropriate independent Gamma distributions given $\mathbf{I}_i$. We then thinned the results by taking every tenth observation and estimated the integrated autocorrelation functions for β and δ, $C_\beta = 1 + 2\sum_{k=0}^{\infty} \text{corr}(\beta_0, \beta_k)$ and $C_\delta = 1 + 2\sum_{k=0}^{\infty} \text{corr}(\delta_0, \delta_k)$, respectively. See Neal and Roberts (2005, Section 5.2), for full details of estimating the integrated autocorrelation function. The estimates of C_β and C_δ were 4.709 and 4.557, respectively, compared with optimal values of 4.291 and 4.088, respectively, reported in Neal and Roberts (2005, Section 5.2). However, our algorithm is quicker to run per iteration since only a single update of $\mathbf{I}$ is required, whereas two update steps are required, the non-centred updating of γ (δ in our notation) and the corresponding updating of infectious periods and the updating of a randomly chosen infectious period for fixed γ. Moreover, the above MCMC algorithm is automatic, whereas the non-centred algorithm requires tuning to find the optimal centring proportion and the proposal variance for the RWM updates of γ. In the case $\alpha = 1$, the estimated posterior means of $\beta = 1.03 \times 10^{-3}$ and $\delta = 0.106$ are similar to those reported in Neal and Roberts (2005) ($\beta = 9.88 \times 10^{-4}$ and $\delta = 0.105$) The posterior mean (standard deviations) for α, δ and β are 33.8(25.2), 2.03(1.40) and $6.34 \times 10^{-4}(1.88 \times 10^{-4})$, respectively, with an estimated posterior mean infectious period of 16.4 days. These results suggest that the general stochastic epidemic model ($\alpha = 1$) is not appropriate for this data.

4. 2001 UK foot and mouth disease data

In this section, we apply the methodology proposed in Section 2 and developed using the homogeneously mixing model in Section 3 to a spatial epidemic outbreak consisting of a subset of the 2001 UK Foot-and-Mouth disease (FMD) outbreak. The FMD outbreak in the UK in 2001 lead to the slaughter of approximately 6 million animals and cost the UK economy billions of pounds, see UK National Audit Office (2002).

The data we consider is the FMD outbreak in Cumbria, a rural county in the North West of England bordering Scotland. The data consists of the geographical location and the number of cattle and sheep on 5378 initially susceptible farms. Cumbria was one of the worst affected counties by the FMD outbreak with 1021 reported cases during the course of the epidemic. A map of the infected and susceptible farms is displayed in Fig. 1. For the 1021 infected farms data is available on the slaughter date, the date upon which all animals on the farm were culled. This data set has previously been analysed by Kypraios (2007) and Jewell et al. (2009) using MCMC. In Kypraios (2007) and Jewell et al. (2009) an SIR model was used with at any given time point, a farm being in one of 3 states, S (Susceptible), I (Infected) and R (Removed/culled). The generic model in Jewell et al. (2009, Section 2), allows for an additional class of notified farm, those farms identified with FMD but not yet culled. Given that there is no information available on notified farms and the general government restrictions placed upon all farms during the FMD outbreak, we follow Jewell et al. (2009, Section 3.2), in making no distinction between infected and notified farms.

Let $\vartheta = (\eta, \zeta, \chi, \xi, \phi)$ and let $\tilde{\vartheta} = (\zeta, \chi, \xi, \phi) = \vartheta_{-\eta}$. We then follow Kypraios (2007) and Jewell et al. (2009) in assuming that the infection rate between an infected farm i and a given farm j is,

$$\lambda_{ij}(\vartheta) = \eta(\zeta(n_i^c)^\chi + (n_i^s)^\chi)(\xi(n_j^c)^\chi + (n_j^s)^\chi)\frac{\phi}{\rho_{ij}^2 + \phi^2} = \eta\tilde{\lambda}_{ij}(\tilde{\vartheta}), \quad \text{say}, \tag{14}$$

where $n_k^c(n_k^s)$ denotes the total number of cattle (sheep) on farm k and ρ_{ij} denotes the Euclidean distance between farms i and j. The reason for focusing upon sheep and cattle numbers on farms is that these were the main drivers of the FMD outbreak, Keeling et al. (2001). Note that η is the baseline infection rate between two farms and ζ and ξ denote the relative infectivity

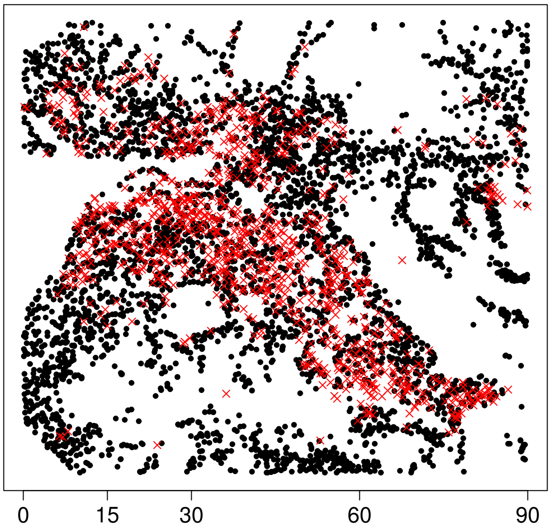

Fig. 1. Susceptible (black solid points) and infected (red cross) farms in Cumbria in km, UK, 2001.

and susceptibility of cattle to sheep, respectively, and χ is a measure of how infectivity and susceptibility of farms vary with increasing size. We would expect $0 \le \chi \le 1$, with $\chi = 0$ denoting no affect of farm size and $\chi = 1$ denoting linear growth in infectivity and susceptibility of farms with increasing size. Finally, ϕ governs the rate of decay in the infection rate with distance between farms. Substituting (14) into (1) gives $\pi(\mathbf{R}, \mathbf{I}|\theta)$, where $\theta = (\vartheta, \alpha, \delta)$. By assigning gamma priors to all the parameters, we obtain the full joint posterior of $\mathbf{I}$ and θ, given by

$$\begin{aligned}\pi(\mathbf{I}, \theta|\mathbf{R}) \propto{}& \prod_{j\neq\kappa}^{m}\left(\sum_{i:I_i<I_j\le R_i}\lambda_{ij}(\vartheta)\right)\exp\left(-\sum_{i=1}^{m}\sum_{j=1}^{m}\lambda_{ij}(\vartheta)(\{R_i\wedge I_j\}-\{I_i\wedge I_j\})-\sum_{i=1}^{m}\sum_{j=m+1}^{N}\lambda_{ij}(\vartheta)(R_i-I_i)\right)\\ &\times\eta^{a_\eta-1}\exp(-\eta b_\eta)\zeta^{a_\zeta-1}\exp(-\zeta b_\zeta)\xi^{a_\xi-1}\exp(-\xi b_\xi)\chi^{a_\chi-1}\exp(-\chi b_\chi)\phi^{a_\phi-1}\exp(-\phi b_\phi)\\ &\times\prod_{i=1}^{m}\left(\frac{\delta^\alpha}{\Gamma(\alpha)}(R_i-I_i)^{\alpha-1}\exp(-\delta(R_i-I_i))\right)\delta^{a_\delta-1}\exp(-\delta b_\delta)\alpha^{a_\alpha-1}\exp(-\alpha b_\alpha).\end{aligned}\tag{15}$$

In the spirit of Liu (1994) and Section 3, it is straightforward to integrate out η and δ. This yields,

$$\begin{aligned}\pi(\mathbf{I}, \vartheta, \alpha|\mathbf{R}) \propto{}& \frac{\Gamma(m+a_\eta)}{(b_\eta+\tilde{A})^{m+a_\eta}}\prod_{j\neq\kappa}^{m}\left(\sum_{i:I_i<I_j\le R_i}\tilde{\lambda}_{ij}(\tilde{\vartheta})\right)\\ &\times\zeta^{a_\zeta-1}\exp(-\zeta b_\zeta)\xi^{a_\xi-1}\exp(-\xi b_\xi)\chi^{a_\chi-1}\exp(-\chi b_\chi)\phi^{a_\phi-1}\exp(-\phi b_\phi)\\ &\times\prod_{i=1}^{m}(R_i-I_i)^{\alpha-1}\frac{\Gamma(\alpha m+a_\delta)}{\Gamma(\alpha)^m(b_\delta+B)^{m\alpha+a_\delta}}\alpha^{a_\alpha-1}\exp(-\alpha b_\alpha),\end{aligned}\tag{16}$$

where $B = \sum_{i=1}^{m}(R_i - I_i)$, the sum of the infectious periods and

$$\tilde{A} = \sum_{i=1}^{m}\sum_{j=1}^{m}\tilde{\lambda}_{ij}(\tilde{\vartheta})(\{R_i\wedge I_j\}-\{I_i\wedge I_j\})+\sum_{i=1}^{m}\sum_{j=m+1}^{N}\tilde{\lambda}_{ij}(\tilde{\vartheta})(R_i-I_i).$$

We now turn to the MCMC algorithm which is based upon (16). At each iteration of the MCMC algorithm, we store $(\tilde{\vartheta}, \alpha, \tilde{A}, B)$ and then samples from δ and η can be obtained, if so desired, from their conditional distributions given $(\tilde{\vartheta}, \alpha, \tilde{A}, B)$ and $\mathbf{R}$. It should be noted that given $\mathbf{I}$ and $\mathbf{R}$, the infection process parameters $\tilde{\vartheta}$ and the infectious period parameter α are independent. Therefore the MCMC algorithm cycles through updating $\mathbf{I}|\alpha, \mathbf{R}, \tilde{\vartheta}$, $\tilde{\vartheta}|\mathbf{R}, \mathbf{I}$ and $\alpha|\mathbf{R}, \mathbf{I}$ at each iteration. As in the homogeneously mixing case, if α is assumed known, omit the update step.

We follow the updating scheme for $\mathbf{I}$ introduced in Section 3.

1. Choose to update p infectious periods with probability u_p and choose $\mathbf{k} = \{k_1, k_2, \dots, k_p\}$ uniformly, at random, from $\{1, 2, \dots, m\}$. We discuss the choice $\mathbf{u}$ shortly.
2. Draw the intermediary γ from the marginal distribution of δ, Gamma$(\alpha m + a_\delta, b_\delta + B)$. Then draw $Q'_{k_1}, Q'_{k_2}, \dots, Q'_{k_p}$ from Gamma (α, γ). Set $I'_{k_j} = R_{k_j} - Q'_{k_j}$ for $j = 1, \dots, p$.

3. The computation of the proposal density $h(\mathbf{Q} \rightarrow \mathbf{Q}')$ and simplifications of the acceptance probability, in terms of cancellation of the infectious period terms, are identical to the homogeneously mixing case in Section 3. Therefore the probability of accepting the proposed move from $\mathbf{I}$ to $\mathbf{I}'$ is

$$\text{Acc} = \min\left\{1, \frac{\prod_{j\neq\kappa'} \sum_{i:I_i'<I_j'\leq R_i} \tilde{\lambda}_{ij}(\tilde{\vartheta})}{\prod_{j\neq\kappa} \sum_{i:I_i<I_j\leq R_i} \tilde{\lambda}_{ij}(\tilde{\vartheta})} \frac{(b_\eta + \tilde{A})^{m+a_\eta}}{(b_\eta + \tilde{A}')^{m+a_\eta}}\right\}.$$

The components of $\tilde{\vartheta}$ were updated componentwise using RWM with the proposal variance updated adaptively using the same updating procedure outlined in (12) and (13) in Section 3. As in Section 3, when α was assumed unknown it was updated using RWM on the logarithmic scale.

For comparison with Jewell et al. (2009), we consider $\alpha = 4$ and α unknown along with $\alpha = 1$ and $\alpha = 20$ for sensitivity analysis. We follow Kypraios (2007) in assigning Gamma(0.001, 0.001) priors to η and δ and Exp(0.001) priors to all the other parameters. Given that $m = 1021$, we choose the initial $\mathbf{u}$ such that $u_j = 1/11$ for $j \in \{1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1021\}$ and $u_j = 0$ otherwise. Then as for the Abakaliki data, we ran the algorithm for 120,000 iterations with the first 20 000 iterations being discarded as burn-in. The first 20 000 iterations split into 2 blocks of 10 000 iterations for running the adaptive procedures for updating $\mathbf{u}$ and the standard deviations of the proposals for the RWM updates. The first 10 000 iterations showed that the acceptance rate was extremely small for $p = 64$ and no proposed moves were accepted for $p > 64$. The optimal performance appears to be for p around 16 and this is used for the second stage proposal for $\mathbf{u}$, where $u_{2j} = 1/13$ for $4 \leq j \leq 16$. It was found that $\tau_p = p\text{Acc}_p$ had a maximum of 3.38 at $p = 12$, although for $p = 10, 12, 14, 16, 18$, $\tau_p > 3.15$ showing consistent results in the range. For the analysis u_{2j} ($4 \leq j \leq 16$) was set proportional to τ_{2j}^3, which meant that approximately 56% of the time we proposed between 10 and 18 infection times to update.

In Fig. 2, the estimated marginal distributions of $\tilde{\vartheta} = (\zeta, \xi, \chi, \phi)$ are shown in the case $\alpha = 4$. These results are indistinguishable from those given in Jewell et al. (2009) with all of the parameters in $\tilde{\vartheta}$ mixing well. Similar estimates of $\tilde{\vartheta} = (\zeta, \xi, \chi, \phi)$ are obtained for the other choices of α with for α unknown, the posterior mean being estimated at 3.76 and the 95% equal tailed credible interval being [2.61, 5.36]. Thus as reported in Jewell et al. (2009), $\alpha = 4$ is well within the support of the distribution. The mis-specification of α either $\alpha = 1$ or $\alpha = 20$ does not have a dramatic effect on the estimation of the parameters $\tilde{\vartheta}$, which suggests that these parameters are robust to mis-specification of the infectious period distribution. The mean infectious period (α/δ) does depend significantly on α rising from 5.4 for $\alpha = 1$ to 9.5 for $\alpha = 20$ and is 7.9 for $\alpha = 4$. However, $\eta \times \alpha/\delta$, the baseline infection rate times the mean infectious period is far more consistent with a mean of 1.4×10^{-6} for $\alpha = 4$ and a mean of 1.7×10^{-6} for $\alpha = 1$ and $\alpha = 20$. Note that $\eta \times \alpha/\delta$ represents the mean total infectious pressure exerted by a farm with a single infected sheep on a farm at the same location consisting of a single susceptible sheep.

We explore the performance of the algorithm using the integrated autocorrelation function as in Section 3. No measures of the mixing of the MCMC algorithms used in Kypraios (2007) and Jewell et al. (2009) are given for comparison and therefore we compare results with those obtained in Section 3 for the homogeneously mixing epidemic model. For comparison we thin the MCMC output using every tenth observation for estimating the integrated autocorrelation function. For fixed $\alpha = 4$, the MCMC algorithm mixes also very well for the infection process parameters, ϑ with the estimated integrated autocorrelation functions being 9.21, 8.54, 15.17, 16.30 and 17.18 for ζ, ξ, χ, ϕ and η, respectively. The mixing of δ is poorer with the estimated integrated autocorrelation functions being 320.28. The reason for this is that a relatively small number of infection times are changed at each iteration but the performance is superior to updating one infection time per iteration. Similar observations are made concerning the mixing of the MCMC algorithm in the case α is unknown.

5. Conclusions

The MCMC algorithm developed in this paper has been designed for partially observed temporal data from an SIR epidemic model. The generic model presented in Section 2 has been demonstrated in this paper for homogeneously mixing (Section 3) and spatially mixing (Section 4) epidemic models. As noted in Section 2, the methodology can easily be applied to household and network epidemic models. The key limitation of the method is the extent to which the independence sampler for the infection periods is an effective mechanism for updating the missing infection times, although this does not appear to be problematic for the cases considered in this paper. The multiple update of infection times is particularly effective when there is considerable uncertainty in who infects who, which is the case in the homogeneously mixing model. The spatial FMD data are more informative about the route of likely infection, and consequently, we find it beneficial to update a smaller proportion of the infection times at a time. It should be straightforward to extend the approach to SEIR models (O'Neill and Becker, 2001), where both the infection times and start of the infectious periods are not observed.

In this paper we found that running the adaptive procedure twice on blocks of 10 000 iterations was sufficient to obtain an efficient MCMC algorithm. It is trivial to adapt the code so that the algorithm could choose how many times the adaptive

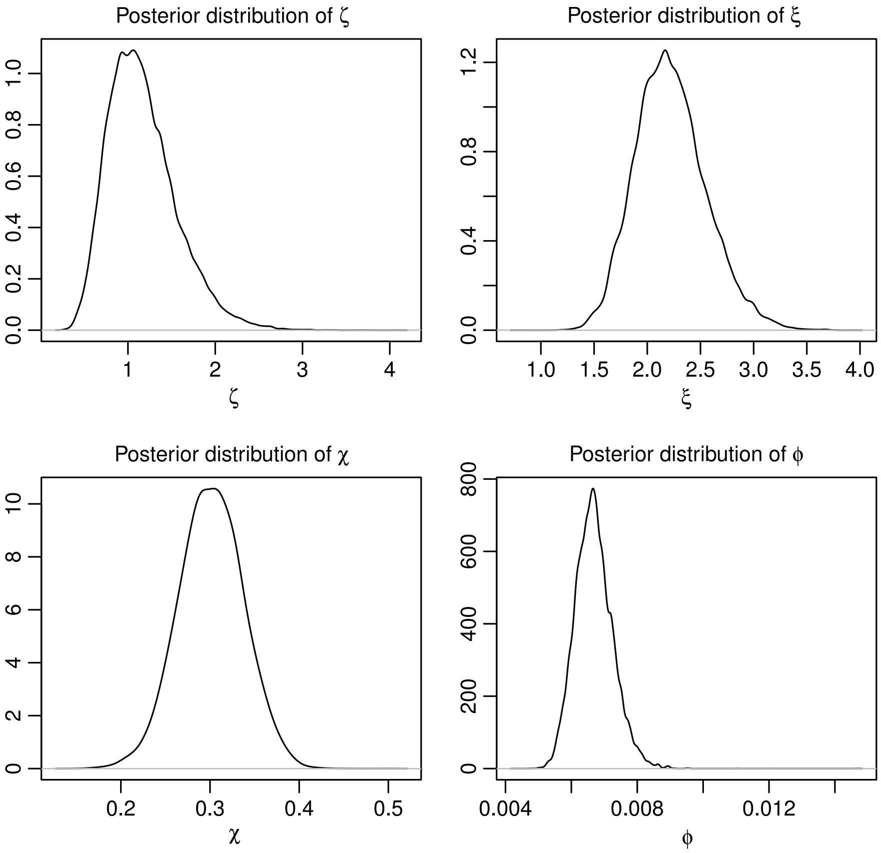

Fig. 2. Estimated posterior distributions of ζ, ξ, χ, ϕ.

procedure was run, stopping when **u** and the proposal standard deviations for the RWM updates do not change significantly between adaptive steps.

Acknowledgements

The authors would like to thank two anonymous referees for their comments and suggestions which have allowed to improve the paper. This research was supported by the Engineering and Physical Sciences Research Council under grant EP/J008443/1.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.csda.2014.07.002.

References

Bailey, N.T.J., 1975. The Mathematical Theory of Infectious Diseases and its Applications, second ed. Griffin, London.
Britton, T., Kypraios, T., O'Neill, P.D., 2011. Inference for epidemics with three levels of mixing: methodology and application to a measles outbreak. Scand. J. Statist. 38, 578–599.
Britton, T., O'Neill, P.D., 2002. Bayesian inference for stochastic epidemics in populations with random social structure. Scand. J. Statist. 29, 375–390.
Gibson, G.J., 1997. Markov chain monte carlo methods for fitting spatiotemporal stochastic models in plant epidemiology. J. R. Stat. Soc. Ser. A 46 (2), 215–233.
Jewell, C.P., Kypraios, T., Neal, P., Roberts, G.O., 2009. Bayesian analysis for emerging infectious diseases. Bayesian Anal. 4 (4), 465–496.
Keeling, M.J., Woolhouse, M.E.J., Shaw, D.J., Matthews, L., Chase-Topping, M., Haydon, D.T., Cornell, S.J., Kappey, J., Wilesmith, J., Grenfell, B.T., 2001. Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. Science 294 (5543), 813–817.
Kypraios, T., 2007. Efficient Bayesian Inference for Partially Observed Stochastic Epidemics and a New Class of Semi-parametric Time Series Models (Ph.D. thesis). Department of Mathematics and Statistics, Lancaster University, Lancaster.
Liu, J.S., 1994. The collapsed Gibbs sampler in Bayesian computations with applications to a gene regulation problem. J. Amer. Statist. Assoc. 89, 958–966.
McKinley, T.J., Ross, J.V., Deardon, R., Cook, A.R., 2014. Simulation-based Bayesian inference for epidemic models. Comput. Statist. Data Anal. 71, 434–447.
Neal, P., Roberts, G.O., 2005. A case study in non-centering for data augmentation: stochastic epidemics. Stat. Comput. 15, 315–327.
O'Neill, P.D., 2009. Bayesian inference for stochastic multitype epidemics in structured populations using sample data. Biostatistics 10 (4), 779–791.
O'Neill, P.D., Becker, N.G., 2001. Inference for an epidemic when susceptibility varies. Biostatistics 2 (1), 99–108.

O'Neill, P.D., Roberts, G.O., 1999. Bayesian inference for partially observed stochastic epidemics. J. R. Stat. Soc. Ser. A 162, 121–129.

Papaspiliopoulos, O., Roberts, G.O., Sköld, M., 2003. Non-centred parametrisations for hierarchical models and data augmentation. In: Bernardo, J.M., Bayarri, M.J., Berger, J.O., Dawid, A.P., Heckerman, D., Smith, A.F.M., West, M. (Eds.), Bayesian Statistics 7. Oxford University Press, New York, pp. 307–326.

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

Roberts, G.O., Rosenthal, J.S., 1998. Optimal scaling of discrete approximations to Langevin diffusions. J. R. Stat. Soc. Ser. B 60, 255–268.

UK National Audit Office, 2002. The 2001 outbreak of foot and mouth disease. Comptroller and auditor general, HC939, Session 2001–2002, The Stationary Office, London.

Xifara, T., 2013. Bayesian Inference on a Coupled Hidden Markov Model for Disease Interactions and a New Position Dependent Metropolis Adjusted Langevin Algorithm (Ph.D. thesis). Department of Mathematics and Statistics, Lancaster University.