Statistica Sinica **20**(2010): Supplement

INFERENCE FOR INDIVIDUAL-LEVEL MODELS OF INFECTIOUS DISEASES IN LARGE POPULATIONS

Rob Deardon[1], Stephen P. Brooks[2], Bryan T. Grenfell[3],
Matthew J. Keeling[4], Michael J. Tildesley[4], Nicholas J. Savill[5],
Darren J. Shaw[5] and Mark E. J. Woolhouse[5]

[1]University of Guelph, [2]University of Cambridge, [3]Penn-State University,
[4]University of Warwick, and [5]University of Edinburgh

Supplementary Material

S1. Imputing the dates of infection of CPs and DCs

The methodology described so far relies on the assumption that the infection status of all farms is known throughout the epidemic. Section 3 detailed how this information was arrived at for IPs. However, no such information is available for CPs and DCs farms since they could have been infected but culled before symptoms appeared. Thus, in order to use the methodology described so far, we must impute the infection status of all culled farms. Under the Bayesian framework already developed here, in which data and parameters are, in effect, interchangeable, it is possible to impute the time of infection of culled farms and, thus, calculate the proportion of culled farms that were infected before slaughter. Unfortunately, we cannot do this for CPs and DCs separately since the data does not allow us confidence in deciding upon whether a substantial proportion farms were culled as a CP or a DC. Due to uncertainty about which culled farms should fall into which category, we treat CPs and DCs as a single category of culls (C). We now consider the estimation of the proportion of these farms which were infected before culling, λ_C. As stated in Section 3, imputing times for other forms of cull, such as the welfare cull, is not considered here.

S1.1. Notation and Model Framework

Let n_C denote the total number of culled (CP and DC) farms, and $C_R(k)$ be the time of cull (removal) for $k = 1, \ldots, n_C$. Further, let $C_E(k)$ denote the number of days before $C_R(k)$ that culled farm k was exposed to FMD. Obviously, $C_E(k) = 0$ implies that no infection of farm k occurred before slaughter and, so, the probability that a culled farm was infected is simply the posterior probability that $C_E(k) > 0$. In addition, we have that $C_E(k) \leq 9$ otherwise, we assume, the farm would have exhibited symptoms and been recorded as an IP.

We now have a set of parameters $C_E(k)$, $k = 1, \ldots, n_C$ which we wish to estimate. We therefore extend (2) to the situation where the parameter vector is now $\boldsymbol{\theta} = \boldsymbol{\theta}_B = (T_s, T_c, S_c, b, \epsilon, \delta_0, k_0, \psi_{S,s}, \psi_{S,c}, \psi_{T,s}, \psi_{T,c}, C_E(1), \ldots, C_E(n_C))$ and the likelihood is given by (2) with $P(i,t) = P_B(i,t)$. The full posterior for these $n_C + 11$ parameters, up to a constant of proportionality, is then explored via MCMC as described below and in Section 7 (results are also given in Section 7).

For ease of exposition we also introduce the following notation. Let $F_{\mathcal{I}}$ denote the set of all farms which at some point in time became an IP with $F_{\mathcal{C}}$ denoting the set of all farms which at some point in time were culled either as a CP or DC. Clearly, $F_{\mathcal{I}} \cap F_{\mathcal{C}} = \emptyset$.

S1.2. Further Reducing Computational Expense

In the 2001 UK epidemic there were 3609 CPs and 1442 DCs (Anderson, 2002), so the imputation of the infection status for these farms could potentially involve the inclusion of an additional 5051 parameters. Even with the techniques described in Section 5, naive implementation of the MCMC algorithm with these additional parameters is prohibitive computationally. In this section we explain how these updates can be performed in a reasonable time by calculating the likelihood ratio corresponding to the current and proposed new infection state for any one farm directly rather than calculating the two likelihoods independently before dividing one by the other.

At any given point in time, t, a culled farm can be in one of four states, $\mathcal{S}, \mathcal{E}, \mathcal{I}, \mathcal{R}$. Whilst the farm is in states $\mathcal{E}$ and $\mathcal{R}$ it makes no contribution to the likelihood since it can neither be infected nor can it infect other farms. Thus, it is the $\mathcal{I}$ and $\mathcal{S}$ states (and the transitions to and from them) that are the key states in terms of the corresponding likelihood contribution. In addition, since $0 \leq C_E(k) \leq 9$ we know that if we update this parameter then the majority of the likelihood contribution for farm k (i.e. before $t = C_R(k) - 9$ and after $C_R(k)$) will remain unchanged. Note that there is no likelihood contribution on the day of cull since we assume that upon cull the farm is neither infectious nor susceptible. Further, in the general case in which we propose updating $C_E(k) = c_1$ to $C_E(k) = c_2$, say, the corresponding likelihood contribution changes only for times $t = C_R(k) - m_c, ..., C_R(k) - 1$, where $m_c = \max(c_1, c_2)$. The change in likelihood is therefore a ratio of products over these times of the likelihood contribution of farm k which, in turn, depends upon the current or proposed new state.

Here, we let $L^{(k)}_{\mathcal{S}}(c)$ denote the contribution to the likelihood of culled farm k, at $t = C^{(k)}_R - c$ if, at time $t = C^{(k)}_R - c + 1$, k is a susceptible farm (i.e. $k \in \mathcal{S}(C^{(k)}_R - c + 1)$). Similarly, we let $L^{(k)}_{\mathcal{E}}(c)$ denote the contribution to the likelihood of culled farm k, at $t = C^{(k)}_R - c$ if, at time $t = C^{(k)}_R - c + 1$, k becomes newly exposed (i.e. $k \in \mathcal{E}(C^{(k)}_R - c + 1) \setminus \mathcal{E}(C^{(k)}_R - c)$). We also let $L^{(k)}_{\mathcal{I}}(c)$ denote the contribution to the likelihood of culled farm k, at $t = C^{(k)}_R - c$ if, at time $t = C^{(k)}_R - c$, k is infectious (i.e. $k \in \mathcal{I}(C^{(k)}_R - c)$). Finally, we let $L^{(k)}_{\circ}(c)$ denote the contribution of culled farm k at time $t = C^{(k)}_R - c$ if farm k is neither susceptible nor newly exposed at time $t = C^{(k)}_R - c + 1$, nor infectious at time $t = C^{(k)}_R - c$.

Clearly,

$$L_{\circ}^{(k)}(c) = 1,$$

$$L_{\mathcal{S}}^{(k)}(c) = 1 - P_B(k, C_R^{(k)} - c),$$

and

$$L_{\mathcal{E}}^{(k)}(c) = P_B(k, C_R^{(k)} - c).$$

Finally, $L_{\mathcal{I}}^{(k)}(c)$ is given by

$$\begin{aligned} L_{\mathcal{I}}^{(k)}(c) &= \left\{ \prod_{i \in \mathcal{S}(t+1)} 1 - P_{B1}(i, C_R^{(k)} - c) \prod_{i \in \mathcal{E}(t+1) \setminus \mathcal{E}(t)} P_{B1}(i, C_R^{(k)} - c) \right\} \\ &\quad \times \left\{ \prod_{i \in \mathcal{S}(t+1)} 1 - P_{B2}^{(k)}(i, C_R^{(k)} - c) \prod_{i \in \mathcal{E}(t+1) \setminus \mathcal{E}(t)} P_{B2}^{(k)}(i, C_R^{(k)} - c) \right\}^{-1} \\ &= \left\{ \prod_{i \in \mathcal{S}(t+1)} 1 - P_{B3}^{(k)}(i, C_R^{(k)} - c) \prod_{i \in \mathcal{E}(t+1) \setminus \mathcal{E}(t)} P_{B1}(i, C_R^{(k)} - c) \right\} \\ &\quad \times \left\{ \prod_{i \in \mathcal{E}(t+1) \setminus \mathcal{E}(t)} P_{B2}^{(k)}(i, C_R^{(k)} - c) \right\}^{-1} \end{aligned} \tag{S1.1}$$

where,

$$P_{B1}(i,t) = 1 - \exp\left\{ -\boldsymbol{S}\mathcal{L}^2(\boldsymbol{N}_i^{\boldsymbol{\psi}_S}; \boldsymbol{\psi}_S, \boldsymbol{\phi}_S) \sum_{j \in \mathcal{I}(t)} \boldsymbol{T}\boldsymbol{N}_j^{\boldsymbol{\psi}_T} K_B(d_{ij}) \right\}$$

$$P_{B2}^{(k)}(i,t) = 1 - \exp\left\{ -\boldsymbol{S}\mathcal{L}^2(\boldsymbol{N}_i^{\boldsymbol{\psi}_S}; \boldsymbol{\psi}_S, \boldsymbol{\phi}_S) \sum_{j \in \mathcal{I}(t) \setminus \{k\}} \boldsymbol{T}\boldsymbol{N}_j^{\boldsymbol{\psi}_T} K_B(d_{ij}) \right\}$$

and

$$P_{B3}^{(k)}(i,t) = 1 - \exp\left\{ -\boldsymbol{S}\mathcal{L}^2(\boldsymbol{N}_i^{\boldsymbol{\psi}_S}; \boldsymbol{\psi}_S, \boldsymbol{\phi}_S) \boldsymbol{T}\boldsymbol{N}_k^{\boldsymbol{\psi}_T} K_B(d_{ik}) \right\}.$$

S1.3. Partitioning the Summations with Culled Farms

The methodology of Section 5 whereby likelihood computation speed was drastically increased by avoiding calculating the entire likelihood every time a parameter was changed as part of the MCMC algorithm, is reliant on the epidemic event history being static over MCMC iterations. Of course, if the infection status of culled farms are allowed to vary, this has implications for the use of this methodology. However, although some increase in computation time is necessary when dealing with variable infection times for culled farms, it is possible to combine the methodology of Section 5 with that of Section S1.2 to produce an algorithm which runs in reasonable time.

The purpose of the work of Section 5 was to avoid having to calculate (7) in its entirety. To extend this this methodology, we rewrite (7) as

$$\chi_B(t) = \prod_{i \in \mathcal{C}_1} \{1 - P_B(i,t)\} \prod_{i \in \mathcal{C}_2} \{(1 - P_B(i,t)\}$$

where $\mathcal{C}_1 = \{i \in \mathcal{S}(t+1) : i \notin F_{\mathcal{C}} \text{ or } i \in F_{\mathcal{C}} \ \& \ t < C_c^{(i)} - 10\}$ and $\mathcal{C}_2 = \{i \in \mathcal{S}(t+1) : i \in F_{\mathcal{C}} \ \& \ t \geq C_c^{(i)} - 10\}$.

It is easy to see that $\prod_{i \in \mathcal{C}_1} \{1 - P_B(i,t)\}$ consists of a product over farms at points in time with a static event history. That is, $\mathcal{C}_1$ consists of all susceptible farms which are not culled, and culled farms for the period up to 10 days before culling (i.e. points in time when a change in C_E does not affect the likelihood). It is therefore possible to use the techniques described in Section 5 to calculate this quantity. The factor $\prod_{i \in \mathcal{C}_2} \{1 - P_B(i,t)\}$ is calculated "in full," as a new likelihood calculation is required depending on the value of C_E at any given time.

The contributions to the likelihood, $L_{\mathcal{E}}^{(k)}(c)$ and $L_{\mathcal{S}}^{(k)}(c)$ are calculated in full each time a new likelihood calculation is required. $L_{\mathcal{I}}^{(k)}(c)$ is treated as above, by splitting the likelihood into parts that have variable event history, and parts that do not. To see this we rewrite (S1.1) as

$$\begin{aligned} L_{\mathcal{I}}^{(k)}(c) \quad = \quad & \left[\prod_{i \in \mathcal{C}_1} \left\{ 1 - P_{B3}(i, C_C^{(k)} - c) \right\} \prod_{i \in \mathcal{C}_2} \left\{ 1 - P_{B3}(i, C_C^{(k)} - c) \right\} \prod_{i \in \mathcal{E}(t+1) \setminus \mathcal{E}(t)} P_{B1}(i, C_C^{(k)} - c) \right] \\ & \times \left[\prod_{i \in \mathcal{E}(t+1) \setminus \mathcal{E}(t)} P_{B2}(i, C_C^{(k)} - c) \right]^{-1}. \end{aligned}$$

Once again, $\prod_{i \in \mathcal{C}_1} \left\{ 1 - P_{B3}(i, C_C^{(k)} - c) \right\}$ can be calculated using the techniques of Section 5, and $\prod_{i \in \mathcal{C}_2} \left\{ 1 - P_{B3}(i, C_C^{(k)} - c) \right\}$ calculated 'in full'.

The above methodology is framed within the context of the model of (6), but this can be easily extended to other models which fall under the general framework of (1).

S2. Model Comparison

The Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002) was used to provide evidence that the newly proposed models here provided an improvement over that of Keeling et al. (2001). The DIC obtained for the model of Keeling et al. (2001) was 1542.11; for the model P_T, 1302.29; and for model P_B, 927.22. These results would therefore imply that P_B provides a better fit to the data than P_B and the original model of Keeling et al. (2001).

S3. Estimation Procedure Simulation Studies

In order to ascertain that the MCMC estimation procedure was working correctly, two simulation studies were carried out. The first of these was used to assess that the parameter estimation is satisfactory in the case where the model is true. To this end, twenty epidemics

were simulated using model P_B and then parameters estimated from this simulated data. The posterior means estimated from the UK 2001 epidemic were used as the parameter values in the simulations (Table 1). The simulations used only IP culling so no cull date imputation was required. (N.B. Some simulations produced no, or very small epidemics, and thus little data. When this occurred the epidemic was re-run in order to produce more informative data.) It was found that posterior means for all estimated parameters were distributed around the true parameter values. For 19 out of the 20 simulated epidemics, the HPDI for all parameters contained the true value. The exceptions to this was one parameterisation which underestimated T_s with a 95% HPDI of $(0.00, 2.24) \times 10^{-1}$ against a true value of 2.62×10^{-1}. Results for a typical (randomly chosen) parmeterisation are shown in Table 2. A similar study showed similar results in terms of parameterising epidemics simulated from model P_T (see Section 7) which used the tracing data kernel of Keeling et al. (2001).

A second simulation study consisted a stochastic ring-culling procedure wherein a fixed proportion, ρ_{RC} of farms within 10km of any IP were culled the day after the IP had been culled. A proportion of farms were culled, as opposed to culling all farms within the given radius since entire ring-culling, even for small rings, tends to produce a very large number of culls. This results in a large number of C_E parameters and leads to very long computation times. $\rho_{RC} = 0.1$ was used (although other values produce similar results). Ten simulations were carried out with the posterior means estimated from the 2001 UK epidemic as parameter values (Table 1). Then ten simulations were carried out with the same parameter values, except T_s and T_c were multiplied by 2.5. These two sets of simulations were used to test the estimation of λ_C for different levels of the true λ_C.

The main parameters of the model were parameterised well, similarly to the first study. The posterior mean estimates of λ_C, along with the true values, are shown in Table 3. As we can see, although there is a large uncertainty in some of the estimates of λ_C, the estimation procedure appears to work reasonably well.

S4. Assessing Control Strategies

Many simulation studies have been undertaken in order to determine the effectiveness of different control strategies for the 2001 UK FMD epidemic (e.g. Keeling et al., 2001, 2003; Tildesley et al., 2006; Ferguson et al., 2001a,b). However, the models used in these papers were parameterised under a classical framework and, hence, uncertainty about the parameters is not fully propagated through to the predictive simulations that underpin the conclusions drawn. Here, we demonstrate how the Bayesian paradigm provides a natural framework for predictive simulation and describe two simulation studies used to explore a ring-culling strategy for the UK epidemic. We base our simulations on the full model of (6) which includes spark infections, non-linearity in the susceptibility and transmissibility of farms, and the change-point kernel. (Note that we are not suggesting that ring-culling would be a useful or efficient FMD control policy. The purpose of the simulation study is purely illustrative of the idea of simulation from the posterior.)

The Bayesian approach to predictive simulation is to produce single epidemics for each of a series of parameter realisations drawn from the posterior distribution. This contrasts with the classical approach in which all simulations are based upon the MLE or some other suitable point estimate. The Bayesian approach therefore incorporates parameter uncertainty into the simulation. For comparison, we compare our results with those obtained by simply taking the posterior mean for each parameter as a fixed basis for all predictive simulations.

Epidemics are simulated in two counties, Cumbria and Devon, which were key to the 2001 epidemic but show different topographic and farm-type distributions. For the purposes of these simulations the two regions are treated in isolation.

Our ring-culling model assumes that IPs are culled 10 days after infection, and the day after an IP is culled, all animals on cattle and/or sheep farms within a radius δ_{rc} of the culled IP are also culled. Our aim is to determine the optimal culling radius. Each simulation begins with the infection of ten farms chosen uniformly from the susceptible set, and ends when there are no more infected farms in the region. One thousand simulations using model P_B were carried out for $\delta_{rc} \in \{0, 250, \ldots, 3000, 4000, \ldots, 10000\}$ in both Cumbria and Devon.

Tables 4 and 5 show results for Cumbria epidemics simulated from the posterior distribution and posterior mean (Table 1) for model P_B. These tables show the mean number of farms on which animals were culled (either as a result of being IP culled or ring-culled), mean number of animals culled, and mean length of epidemic, over all simulated epidemics, for ring-culls up to 3000m in radius. Tables 6 and 7 show similar results for Devon.

Figure 6 shows plots for the number of farms on which animals were culled, number of animals slaughtered, and length of epidemic against δ_{rc}, for the both the simulations based upon draws from the posterior density, and those which used the posterior mean as a basis for the simulations, in Cumbria. Figure 7 provides similar plots for Devon. In Figures 6 and 7 mean results for each δ_{rc} are shown as a continuous line, and individual simulation results are shown as points.

Perhaps the first thing to notice when comparing Tables 4 and 5, and, Tables 6 and 7, is that the posterior mean simulations tend to produce much more intense epidemics than the posterior sampled simulations; they result in more farms being affected, more animals being culled, and longer epidemics. The differences between the posterior mean and posterior simulations are larger in situations when less culling is carried out.

In terms of optimising outcomes, it is clear, whether we consider the posterior mean, or posterior sampled, simulations, the larger the ring-cull, the shorter the epidemic. For example, in the posterior mean Cumbria study, length of epidemic is reduced from an average of 107.1 days with no ring-cull to 27.5 days with a 3km ring-cull; in the posterior sampled Cumbria study, the reduction is from 77.6 days to 24.8 days. Such an outcome may be of interest if we wish to minimise the effect of a foot-and-mouth disease outbreak on the whole UK economy. It is plain to see, however, that maximising the ring-cull does not optimise our other two criteria.

In terms of these two outcomes, basing the two simulations on the posterior mean and ignoring parameter uncertainty appears to give quite misleading results. For example, in the

case of Cumbria using the posterior mean simulation study, the optimal ring size appears to be 1500m, both in terms of minimising affected farms, and minimising numbers of animals culled. This also results in a sizable improvement over no ring-culling (an average of 178.7 farms affected with 1500m ring-cull against 261.4 farms affected with no ring-cull; and 86391.2 animals culled with 1500m ring-cull against 263585.7 animals culled with no ring-cull). However, in the posterior sampled simulations we see far fewer farms affected and animals culled, for given ring-cull sizes, than for the posterior mean simulations. For example, in the posterior mean simulations with no ring-cull we see an average of 106.3 farms affected by the epidemic (261.4 in the posterior sampled simulations) and 103968.1 animals culled on average (263586.7 in the posterior sampled simulations). Also, if posterior sampling is used then ring-culling tends to increase the number of farms on which animals are culled (there is a slight reduction in the mean number of farms culled against no ring-culling if a ring-cull of 750m is used; 104.7 against 106.3), suggesting that perhaps no ring-culling should be carried out if we consider this our primary response. If we wish to minimise the number of animals culled, these simulations suggest a ring-cull of 1250m (similar to the posterior mean simulations). However, the improvement is not as drastic in terms of animals saved as for the posterior sampled simulations (55463.8 with 1250m ring-cull against 103968.1 with no ring-cull) and, of course, this strategy has a detrimental effect in terms of the total number of farms on which animals are culled (120.9 with 1250m ring-cull against 106.3 with no ring-cull).

Figure 6 shows that, in the posterior mean simulation with no (or small scale ring-culling) both the distributions for the number of farms affected, and the numbers of animals culled, are bimodal. This is, of course, a common trait in stochastic models (see for example, Andersson and Britton, 2000). Consider, for example, the distribution of the total number of farms infected and/or culled over the 1000 simulations in which there was no ring-culling for the posterior mean and posterior sampled Cumbria simulations, shown in Figure 8. We can see that, in the case of the posterior mean simulations, one mode of this distribution is centred around very small scale epidemics of around 10-50, and the other mode is centred around larger epidemics of around 800 - 950 farms. Ring-culling appears to help by making it more likely that the potentially large epidemics are stamped out early.

In the posterior sampled simulations, however, the distribution of farms affected by the epidemic (and distribution of numbers of animals culled) when there is no ring-culling, although still bimodal, is less polarised. That is, there is a mode around epidemics of size 10-50 (ten being the initial number of farms infected) and a mode around 500-600. However, the distribution has visible mass between these two modes and spreads out up to around epidemic size 950. (This may give the strange initial impression that the posterior sampled simulations actually have less variation than the mean sampled simulations. This is, of course, not true but due to the fact that the extra variation is realised in a way that produces more mass between the two modes. Furthermore, since we have only a finite number of simulations, in the case of the posterior sampled simulations, the mass existing between the two (flatter) modes means the upper tail is less protracted. If a larger number of posterior sampled, as well as posterior mean

simulations, without ring-culling are run, then the distribution of the number of farms affected by the epidemic does visibly encompass that of the distribution obtained from the posterior mean simulations.) Thus, if we consider the parameter uncertainty inherent in the posterior sampled simulations, we conclude that the consequences of having no ring-cull are less serious than in the posterior mean simulations.

In Devon, the differences between posterior sampled, and posterior mean, simulation studies are less marked. Results for the posterior mean simulations (Table 6) imply that a small improvement can be made in the number of animals culled by using a ring-cull of size 750m (102015.2 animals), as opposed to no ring-cull (123268.4 animals). However, this improvement is small compared to either the posterior sampled, or indeed posterior mean, results in Cumbria. There appears to be no improvement to be made on no ring-culling for the numbers of farms affected (175.2).

In the case of the posterior sampled simulations for Devon, we see much smaller epidemics (for simulations with no ring-cull, an average of 67.3 farms affected and 44863.8 animals culled) than in the case of the posterior mean simulations. This implies, once again, that taking account of parameter uncertainty, at least in this situation, should drastically reduce our fear about the likely damage of an outbreak. In the case of the posterior sampled simulations for Devon, once again, there appears to be no improvement to be made from any sort of ring-cull.

One general conclusion here seems to be that taking into account parameter uncertainty when devising control strategy can drastically alter our perceptions about potential epidemic intensity, and it is possible, in certain circumstances, that this may result in a change in recommended culling policy. (Another possible conclusion is that optimal control policy may vary quite widely over different regions. For example, in the case of the posterior simulated simulations, if we are primarily interested in minimising the numbers of animals culled, Devon and Cumbria appear to demand quite different policies.) Of course, as already stated, we do not present this study as part of any prescription for an actual control policy, but to show how the use of distributional, rather than point, parameter estimates can be used to explore epidemic and control strategy dynamics.

Appendix: Expressions for ζ terms of Section 5.

$\zeta^{(q)}_{1,x}$, for $x \in \{1,2,3\}$ and $q \in \{s,c\}$ for use in (8)

$$\begin{aligned}
\zeta^{(q)}_{1,1}(t) &= \sum_{i \in \mathcal{S}(t+1)} {N_{s,i}}^{\phi_{S,s}} \\
\zeta^{(q)}_{1,2}(t) &= \sum_{i \in \mathcal{S}(t+1)} {N_{s,i}}^{\phi_{S,s}} \ln(N_{s,i}) \\
\zeta^{(q)}_{1,3}(t) &= \sum_{i \in \mathcal{S}(t+1)} \frac{1}{2} {N_{s,i}}^{\phi_{S,s}} \ln^2(N_{s,i}) \\
\zeta^{(q)}_{2,1}(t) &= \sum_{i \in \mathcal{S}(t+1)} {N_{c,i}}^{\phi_{S,c}} \\
\zeta^{(q)}_{2,2}(t) &= \sum_{i \in \mathcal{S}(t+1)} {N_{c,i}}^{\phi_{S,c}} \ln(N_{c,i}) \\
\zeta^{(q)}_{2,3}(t) &= \sum_{i \in \mathcal{S}(t+1)} \frac{1}{2} {N_{c,i}}^{\phi_{S,c}} \ln^2(N_{c,i})
\end{aligned}$$

$\zeta^{(q)}_{2,x}(j,t)$, for $x \in \{1,\ldots 9\}$ and $q \in \{s,c\}$ for use in (10)

$$\begin{aligned}
\zeta^{(q)}_{2,1}(j,t) &= \sum_{i \in \mathcal{S}(t+1)} d_{ij} {N_{q,i}}^{\psi_{S,q}} \\
\zeta^{(q)}_{2,2}(j,t) &= \sum_{i \in \mathcal{S}(t+1)} d_{ij} {N_{q,i}}^{\psi_{S,q}} \ln(d_{ij}) \\
\zeta^{(q)}_{2,3}(j,t) &= \sum_{i \in \mathcal{S}(t+1)} \frac{1}{2} d_{ij} {N_{s,i}}^{\psi_{S,q}} \ln^2(d_{ij}) \\
\zeta^{(q)}_{2,4}(j,t) &= \sum_{i \in \mathcal{S}(t+1)} d_{ij} {N_{q,i}}^{\psi_{S,q}} \ln(N_{s,i}) \\
\zeta^{(q)}_{2,5}(j,t) &= \sum_{i \in \mathcal{S}(t+1)} d_{ij} {N_{q,i}}^{\psi_{S,q}} \ln(d_{ij}) \ln(N_{s,i}) \\
\zeta^{(q)}_{2,6}(j,t) &= \sum_{i \in \mathcal{S}(t+1)} \frac{1}{2} d_{ij} {N_{s,i}}^{\psi_{S,q}} \ln^2(d_{ij}) \ln(N_{s,i}) \\
\zeta^{(q)}_{2,7}(j,t) &= \sum_{i \in \mathcal{S}(t+1)} d_{ij} {N_{q,i}}^{\psi_{S,q}} \ln^2(N_{s,i}) \\
\zeta^{(q)}_{2,8}(j,t) &= \sum_{i \in \mathcal{S}(t+1)} d_{ij} {N_{q,i}}^{\psi_{S,q}} \ln(d_{ij}) \ln^2(N_{s,i}) \\
\zeta^{(q)}_{2,9}(j,t) &= \sum_{i \in \mathcal{S}(t+1)} \frac{1}{2} d_{ij} {N_{q,i}}^{\psi_{S,q}} \ln^2(d_{ij}) \ln^2(N_{s,i})
\end{aligned}$$

References

Anderson, I. (2002). *Foot and mouth disease 2001: Lessons to be learned inquiry*. The Stationary Office, London.

Andersson, H. and Britton, T. (2000). *Stochastic Epidemic Models and their Statistical Analysis*. Number 151 in Lecture Notes in Statistics. Springer-Verlag, New York.

Ferguson, N. M., Donnelly, C. A., and Anderson, R. M. (2001a). The foot-and-mouth epidemic in Great Britain: pattern of spread and impact of interventions. *Science*, 292:1155–1160.

Ferguson, N. M., Donnelly, C. A., and Anderson, R. M. (2001b). Transmission intensity and impact of control policies on the foot-and-mouth epidemic in Great Britain. *Nature*, 413:542–548.

Keeling, M. J., Woolhouse, M. E. J., May, R. M., Davies, G., and Grenfell, B. T. (2003). Modelling vaccination strategies against foot-and-mouth disease. *Nature*, 421:136–142.

Keeling, M. J., Woolhouse, M. E. J., Shaw, D. J., Matthews, L., Chase-Topping, M., Haydon, D., Cornell, S. J., Kappey, J., Wilesmith, J., and Grenfell, B. (2001). Dynamics of UK foot-and-mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science*, 294:813–817.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B*, 64(4):583–639.

Tildesley, M. J., Savill, N. J., Shaw, D. J., Deardon, R., Brooks, S. P., Woolhouse, M. E. J., Grenfell, B. T., and Keeling, M. J. (2006). Optimal reactive vaccination strategies for an outbreak of foot-and-mouth disease in Great Britain. *Nature*, 440:83–86.

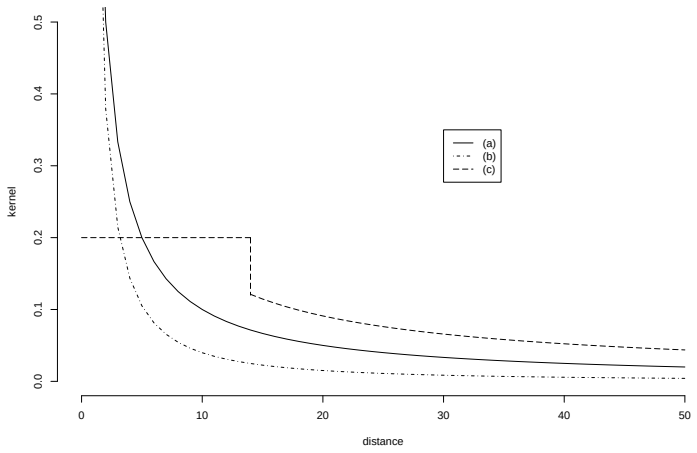

Figure 4: Example change-point distance kernels, $K_A(d_{ij})$: (a) $b = -1.0$, $\delta_0 = 0$; (b) $b = -1.4$, $\delta_0 = 0$; and (c) $b = -0.8$, $\delta_0 = 15$; $k_0 = 0.2$

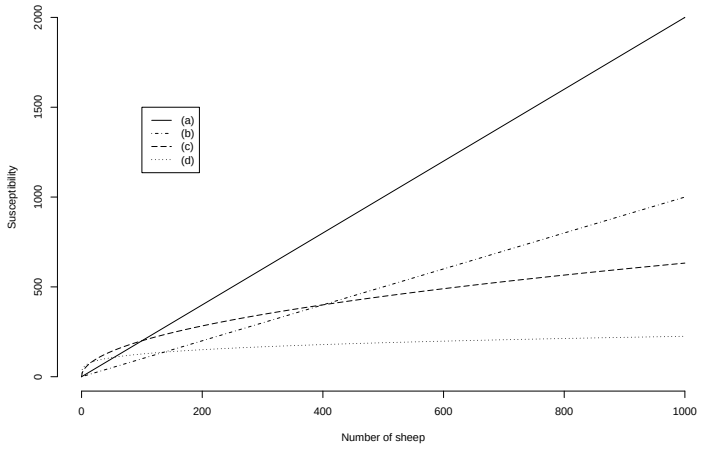

Figure 5: Example susceptibility functions against number of sheep, ${S_s N_{i,s}}^{\psi_{S,s}}$:
(a) $S_s = 2.0$, $\psi_{S,s} = 1.0$; (b) $S_s = 1.0$, $\psi_{S,s} = 1.0$; (c) $S_s = 20.0$, $\psi_{S,s} = 0.5$; and
(d) $S_s = 40.0$, $\psi_{S,s} = 0.25$. N.B. The same examples hold for equivalent sheep/cattle and/or susceptibility/transmissibility functions.

Table 2: Statistics of marginal posterior densities for simulation study under P_B

Parameter	**Posterior Mean Estimate (95% HPDIs)**	**True Value**
b	-1.67 (-1.73, -1.59)	-1.66
T_s	2.09×10^{-1} $(0.01, 4.86)\times10^{-1}$	2.62×10^{-1}
T_c	1.65×10^{-1} $(0.01, 5.16)\times10^{-1}$	1.22×10^{-1}
S_s	1.00	1.00
S_c	7.23 (2.01, 12.2)	7.14
ϵ	-2.05×10^{-10} $(-60.1, 0.00)\times10^{-10}$	-2.45×10^{-10}
k_0	1.39×10^{-5} $(0.41, 2.46)\times10^{-5}$	1.85×10^{-5}
δ_0	1038 (595,1565)	719
$\psi_{T,s}$	0.14 (0.000, 0.30)	0.074
$\psi_{T,c}$	0.36 (0.000, 0.79)	0.32
$\psi_{S,s}$	0.89 (0.81, 0.99)	0.91
$\psi_{S,c}$	0.90 (0.74, 1.07)	0.87

Table 3: Results of simulation study to test estimation of λ_C

UK 2001 epidemic parameters		**UK 2001 epidemic parameters but with increased transmissibility**	
Posterior Mean Estimate (95% HPDIs)	**True Value**	**Posterior Mean Estimate (95% HPDIs)**	**True Value**
0.066 (0.000, 0.121)	0.036	0.184 (0.072, 0.260)	0.135
0.058 (0.000, 0.128)	0.047	0.142 (0.058, 0.231)	0.158
0.032 (0.000, 0.144)	0.048	0.161 (0.089, 0.264)	0.175
0.031 (0.000, 0.095)	0.068	0.158 (0.101, 0.257)	0.230
0.039 (0.000, 0.136)	0.030	0.223 (0.093, 0.291)	0.176
0.024 (0.000, 0.091)	0.034	0.173 (0.065, 0.282)	0.187
0.029 (0.000, 0.086)	0.031	0.172 (0.103, 0.289)	0.189
0.056 (0.000, 0.076)	0.073	0.159 (0.098, 0.239)	0.183
0.029 (0.000, 0.087)	0.020	0.163 (0.077, 0.241)	0.147
0.041 (0.000, 0.104)	0.049	0.173 (0.115, 0.276)	0.222

Table 4: Statistics of interest for Cumbria posterior mean simulation study

Distance (m)	**Mean no. of culled farms**	**Mean no. of culled animals**	**Mean end date of epidemic**
0.0	261.4	263585.7	107.1
250.0	273.7	253063.4	100.8
500.0	271.1	211423.4	91.9
750.0	251.5	169931.7	76.7
1000.0	223.6	133016.8	63.3
1250.0	198.4	105190.3	51.7
1500.0	178.7	86391.2	43.3
1750.0	193.5	89366.8	38.7
2000.0	212.6	92116.4	35.6
2250.0	226.1	97712.2	33.1
2500.0	237.4	101159.0	29.4
2750.0	260.1	112127.8	28.9
3000.0	284.1	118662.5	27.5

Table 5: Statistics of interest for Cumbria posterior sampled simulation study

Distance (m)	**Mean no. of culled farms**	**Mean no. of culled animals**	**Mean end date of epidemic**
0.0	106.3	103968.1	77.6
250.0	118.0	105975.1	77.6
500.0	124.0	90692.8	67.0
750.0	104.7	62945.7	52.3
1000.0	111.5	58122.6	43.7
1250.0	120.9	55462.8	38.9
1500.0	132.7	59500.3	36.3
1750.0	148.9	63747.6	31.8
2000.0	178.5	73332.5	30.8
2250.0	197.1	81218.1	28.2
2500.0	218.8	90878.6	28.4
2750.0	240.7	100121.9	26.4
3000.0	261.8	106444.3	24.8

Table 6: Statistics of interest for Devon posterior mean simulation study

Distance (m)	**Mean no. of culled farms**	**Mean no. of culled animals**	**Mean end date of epidemic**
0.0	175.2	123268.4	147.4
250.0	198.4	116937.5	132.7
500.0	228.9	112825.0	120.1
750.0	254.8	102105.2	97.7
1000.0	292.4	108255.0	83.2
1250.0	328.4	114434.2	69.4
1500.0	375.8	119314.9	60.5
1750.0	451.1	141117.6	56.5
2000.0	527.3	159920.1	53.3
2250.0	587.9	170259.6	48.9
2500.0	647.7	182977.1	46.1
2750.0	712.7	199241.0	43.4
3000.0	819.3	229332.9	42.7

Table 7: Statistics of interest for Devon posterior sampled simulation study

Distance (m)	**Mean no. of culled farms**	**Mean no. of culled animals**	**Mean end date of epidemic**
0.0	67.3	44863.8	85.6
250.0	83.6	46087.5	83.4
500.0	103.4	46570.1	77.5
750.0	150.6	56231.8	71.0
1000.0	183.7	66092.7	62.1
1250.0	229.6	77961.0	53.9
1500.0	295.3	91070.0	50.5
1750.0	348.1	106075.8	45.7
2000.0	419.6	123564.6	44.0
2250.0	494.4	140347.6	42.8
2500.0	562.1	155478.7	40.8
2750.0	634.7	174292.3	38.5
3000.0	713.9	194370.1	37.8

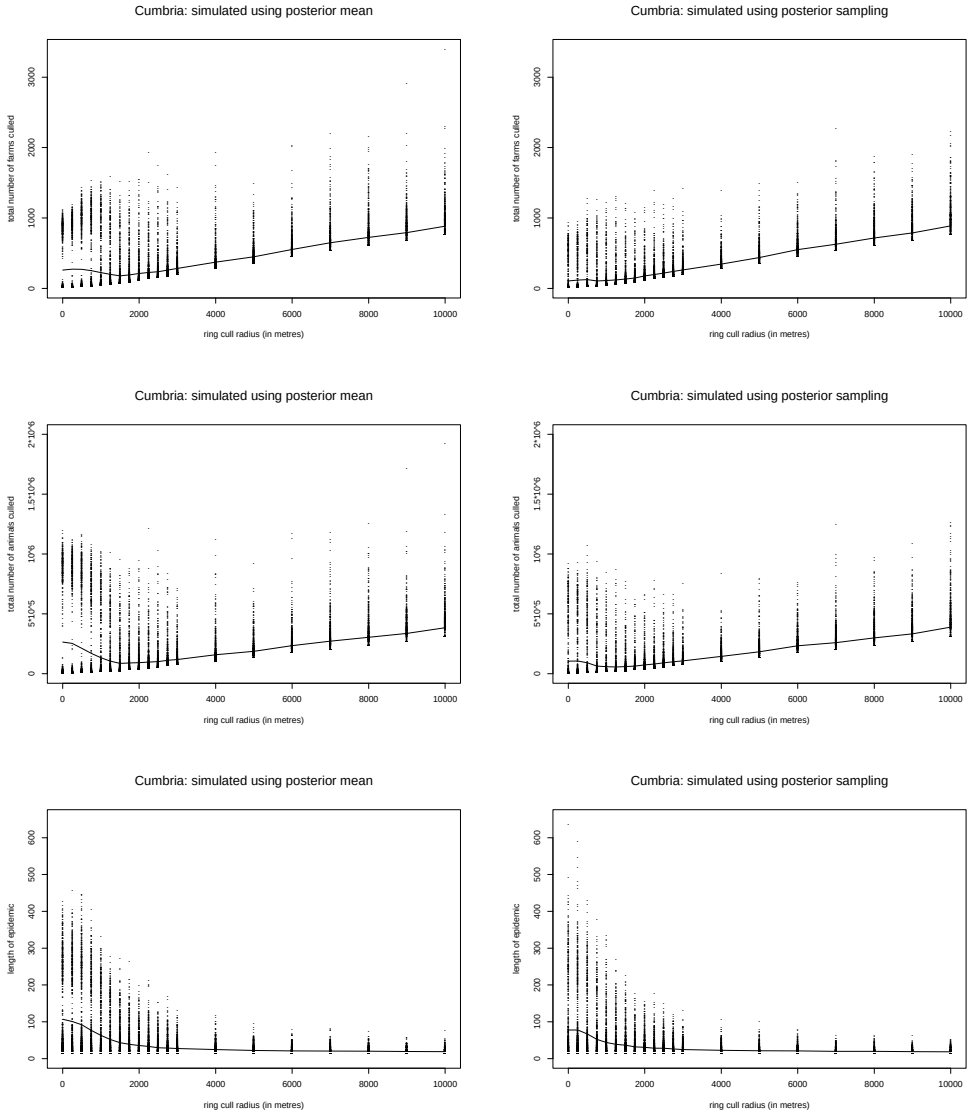

Figure 6: Results for Cumbria Simulation Study

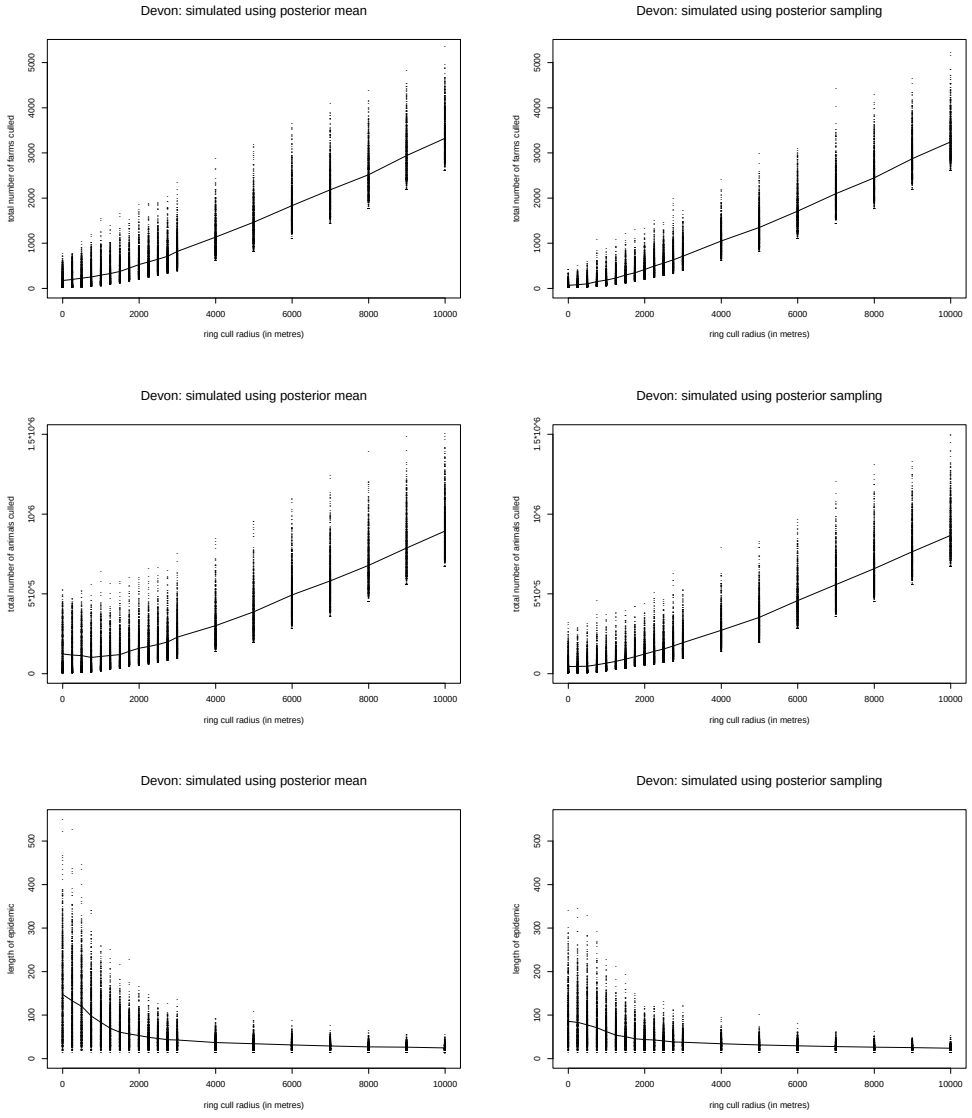

Figure 7: Results for Devon Simulation Study

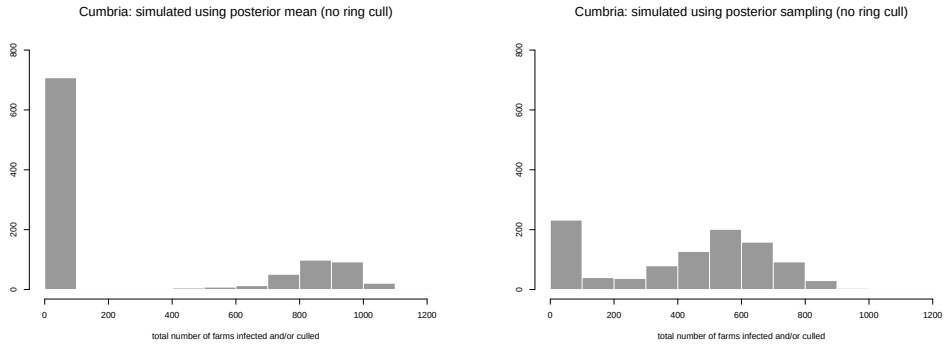

Figure 8: Final epidemic size distributions under no ring culling for Cumbria simulation Study