STATISTICS IN MEDICINE
Statist. Med. 2002; **21**:2703–2721 (DOI: 10.1002/sim.1217)

Hierarchical statistical modelling of influenza epidemic dynamics in space and time

Andrew S. Mugglin[1,*,†], Noel Cressie[2] and Islay Gemmell[3]

[1] *Arrhythmia Management Clinical & Outcomes Research, Medtronic Inc., 7000 Central Avenue NE, MS CW300, Minneapolis MN 55432-3576, U.S.A.*
[2] *Department of Statistics, The Ohio State University, 1958 Neil Avenue, Columbus OH 43210, U.S.A.*
[3] *Drug Misuse Research Unit, University of Manchester, Bury New Road, Manchester M25 3BL, U.K.*

SUMMARY

An infectious disease typically spreads via contact between infected and susceptible individuals. Since the small-scale movements and contacts between people are generally not recorded, available data regarding infectious disease are often aggregations in space and time, yielding small-area counts of the number infected during successive, regular time intervals. In this paper, we develop a spatially descriptive, temporally dynamic hierarchical model to be fitted to such data. Disease counts are viewed as a realization from an underlying multivariate autoregressive process, where the relative risk of infection incorporates the space–time dynamic. We take a Bayesian approach, using Markov chain Monte Carlo to compute posterior estimates of all parameters of interest. We apply the methodology to an influenza epidemic in Scotland during the years 1989–1990. Copyright © 2002 John Wiley & Sons, Ltd.

KEY WORDS: small-area counts; multivariate autoregression; Markov chain Monte Carlo; Markov random field; infectious disease

1. INTRODUCTION

For reasons of confidentiality or practicality, disease-incidence data are often reported as small-area counts or rates over a series of time periods. When the disease is non-infectious, the time period is typically on the order of one or more years. Here, interest centres on covariates as explanatory variables, so that the disease aetiology might be understood and public policy might be altered to ameliorate the effects of the disease. The part of the model that incorporates spatial and temporal statistical dependence often serves the secondary role

*Correspondence to: Andrew S. Mugglin, Arrhythmia Management Clinical & Outcomes Research, Medtronic Inc., 7000 Central Avenue NE, MS CW300, Minneapolis, MN 55432-3576, U.S.A.
†E-mail: andrew.mugglin@medtronic.com

Contract/grant sponsor: U.S. Environmental Protection Agency; contract/grant number: R827-257-01-0
Contract/grant sponsor: Office of Naval Research; contract/grant number: N00014-99-1-0001

Received September 2000
Accepted December 2001

of absorbing deviations from the explanatory part. When the disease is infectious, however, the time period is typically of much smaller duration, often daily or weekly. Here, interest focuses on the space–time spread of the disease, and covariates can simply serve to stratify the expected incidence or prevalence.

The most common models used in analysing infectious diseases are the so-called SEIR (susceptible/exposed/infectious/removed) models, which have been extensively studied in the mathematical biology literature (for example, Anderson and May [1]). These are a type of compartmental model (Jacquez [2]) where individuals move from one compartment or state to another. The local behaviour of individuals leads to a set of mean-field partial differential equations. Until recently, most models for the behaviour of infectious diseases and epidemic spread focused on theoretical stochastic models, usually confined to temporal dynamics and demographic heterogeneities. In recent years, there have been some efforts to extend these models to include correlation of rates of change between neighbouring sites or individuals; see, for example, Bolker and Grenfell [3] or Keeling *et al.* [4]. Others have focused their interest on the detection of potential epidemics from early warning signs using cumulative sum techniques (for example, O'Brien and Christie [5]).

Our goal in this paper is to take a 'field' view rather than a 'particle' view of the space–time spread of an epidemic, incorporating in our model much of the sophisticated Bayesian hierarchical technology used for modelling non-infectious diseases (for example, Waller *et al.* [6], Knorr-Held and Besag [7], Cressie *et al.* [8] and Lawson and Leimich [9]). Our approach is to capture diffusion as well as growth and recession of the disease through a spatially descriptive, temporally dynamic hierarchical model. As we indicated above, covariates serve as stratifying variables, but our main interest is in the space–time dynamics of the model. As a proof of concept, Cressie and Mugglin [10] present a similar model that they fit to *simulated* disease-count data, simulated from the very model they are fitting. They are able to recover the true parameters, assess goodness-of-fit, and predict the underlying 'hidden' process successfully. In contrast to this hierarchical approach that attempts to model each potential source of variability, Carrat and Valleron [11] present a geostatistical study of an outbreak of an influenza-like illness. Their spatial model implicitly assumes that rates have homogeneous variances, whereas our model recognizes that rates based on small populations are more variable than those based on large populations.

The attractiveness of a Bayesian hierarchical approach is that an extremely complicated model can be built out of a succession of relatively simple components. Variability is accounted for in a straightforward way, and the resulting posterior distributions can be used to answer many different questions of statistical inference. Historically, computing the posterior has been intractable except in certain special cases, but the advent of the numerical technique known as Markov chain Monte Carlo (for example, Gilks *et al.* [12]) has made simulation from the posterior possible and has opened wide the door for practical Bayesian analyses.

The remainder of this paper is organized as follows. In Section 2, we develop a Bayesian hierarchical statistical model for infectious-disease counts in space and time. Then, in Section 3, we apply our model to a data set involving an influenza epidemic in Scotland in the years 1989–1990. After fitting the model, we interpret the results through the posterior distributions of model parameters, relative risks, and a number of derived quantities such as velocity of disease growth. Finally, in Section 4, we discuss both the future potential for, and the limitations of, our model.

2. DYNAMIC SPATIAL MODEL OF SMALL-AREA DISEASE COUNTS

To illustrate our approach to modelling epidemics, we begin with a simple, univariate process that attempts to model growth and decline after a period of stability. Consider the univariate first-order autoregressive (AR(1)) Gaussian time series $\{s_t\colon\ t=1,2,\ldots\}$ that behaves according to the equation

$$s_t = \eta s_{t-1} + \varepsilon_t \tag{1}$$

where η is a fixed constant, $\{\varepsilon_t\}$ is a sequence of independent and identically distributed Gaussian(β,σ^2) random variables, and ε_t and s_{t-1} are independent of each other. Stationarity implies that $|\eta|<1$, and $\{s_t\}$ has a mean value of $\beta/(1-\eta)$, with variance $\sigma^2/(1-\eta^2)$ and $\mathrm{corr}(s_t, s_{t-h}) = \eta^h$.

In modelling an infectious-disease epidemic, we think of a time series that, for a while, exhibits non-stationary growth before declining again to vary around some stationary mean. However, choosing $|\eta|>1$ does not allow us to turn on the epidemic's growth in a controlled way. For example, if we choose $\eta>1$, then $\mathrm{var}(s_t) = \eta^2\mathrm{var}(s_{t-1}) + \mathrm{var}(\varepsilon_t)$, which implies that the variance is growing. Furthermore, $\mathrm{corr}(s_t, s_{t-1}) = \{1-\mathrm{var}(\varepsilon_t)/\mathrm{var}(s_t)\}^{1/2}$, which approaches unity as time progresses. Additionally, a moderate positive value of ηs_{t-1} could be dominated by a negative realization ε_t, and the series could then assume a negative trajectory that, once started, would be difficult to stop. The instability is even worse if $\eta<-1$.

We seek a method of turning an epidemic 'on' and 'off' in a more controlled way. Consider instead the autoregressive time series (1), where the mean of $\{\varepsilon_t\}$ now varies. Suppose $\beta=\beta_0$ for $t<t_0$ and $\beta=\beta_1>\beta_0$ for $t\geqslant t_0$; for $t\geqslant t_0$, the series will tend to grow until a new stationary mean $E(s_t)=\beta_1/(1-\eta)$ is reached. A decline back to $\beta_0/(1-\eta)$ can be achieved by specifying $\beta=\beta_0$ again for $t\geqslant t_1>t_0$. In this manner, we can build a piecewise constant sequence of β's that allows $\{s_t\}$ to change into a growth phase and a subsequent recession phase, with possible intermediate 'stopping points' along the way. For instance, we might select $t_2>t_1>t_0$ and $\beta_1>\beta_2>\beta_0$, and set the mean of $\{\varepsilon_t\}$ in (1) to be

$$\beta_{\rho(t)} \equiv \begin{cases} \beta_0 & \text{if } t<t_0, \quad \text{for stability} \\ \beta_1 & \text{if } t_0\leqslant t<t_1, \quad \text{for growth} \\ \beta_2 & \text{if } t_1\leqslant t<t_2, \quad \text{for intermediate decline} \\ \beta_0 & \text{if } t\geqslant t_2, \quad \text{for final decline to stability} \end{cases} \tag{2}$$

Now we return to the spatio–temporal model for infectious-disease counts and generalize (1) to a multivariate (spatial) model, 'hidden' behind Poisson counts. Consider disease counts y_{it} in region i $(i=1,\ldots,I)$ during time period t $(t=1,\ldots,T)$, where conditionally and independently

$$y_{it}\,|\,z_{it} \sim \mathrm{Poisson}(E_i e^{z_{it}}) \tag{3}$$

In other words, the counts are Poisson-distributed with mean parameter $E_i e^{z_{it}}$. Here E_i is the number of cases expected to occur in region i in a unit time interval under non-epidemic conditions (see Section 3 for discussion of its estimation), and z_{it} is the logarithm of the

relative risk that accounts for a departure from the expected number of cases. It is on this scale that we model spatial and temporal dependence through a generalization of (1).

If we have spatially or temporally varying explanatory variables, we can include them in the model by combining them into a vector $\mathbf{x}_{it}$ and writing

$$z_{it} = \mathbf{x}_{it}'\boldsymbol{\alpha} + s_{it} \tag{4}$$

where $\boldsymbol{\alpha}$ is a vector of regression coefficients. We do *not* include an intercept term in $\mathbf{x}_{it}'\boldsymbol{\alpha}$ because we choose to associate an intercept with s_{it} and use it as in (2) to turn the epidemic on and off. Writing $(s_{1t}, s_{2t}, \ldots, s_{It})'$ as $\mathbf{s}_t$, define the multivariate Gaussian AR(1) process through

$$\mathbf{s}_t = H\mathbf{s}_{t-1} + \boldsymbol{\varepsilon}_t \tag{5}$$

where H is an $I \times I$ autoregressive coefficient matrix, and $\boldsymbol{\varepsilon}_t$ (independent of $\mathbf{s}_{t-1}$) is the epidemic-forcing term that we assume to be a realization from a Gaussian Markov random field. Specifically, we assume

$$\boldsymbol{\varepsilon}_t \sim \mathrm{MVN}(\beta_{\rho(t)}\mathbf{1}, \Sigma) \tag{6}$$

where MVN(.,.) denotes the multivariate normal distribution, $\mathbf{1}$ is a vector of ones, $\rho(t) = 0, 1,$ or 2, indicating the stage of the disease at time t ($0 =$ stable, no epidemic; $1 =$ growth; $2 =$ recession to some intermediate level prior to receding to stability) as given by (2), and Σ is the variance–covariance matrix for a Markov random field. The model assumes that the mean forcing is global, without spatial heterogeneity. Spatial dependence is included in the structures of the matrices H and Σ, leading to the interaction between temporal and spatial components given by the autoregression (5). The general form of Σ is

$$\Sigma = \sigma^2(I - \phi C)^{-1}M \tag{7}$$

where M is a diagonal matrix of conditional variances, C is a matrix of partial-regression coefficients with zeros down the diagonal, and ϕ is a spatial dependence parameter such that Σ is positive definite. This family of models is contained in the class of conditional autoregressive (CAR) models; see Cressie, reference [13], Section 6.4. Following Cressie and Chan [14], M has entries E_i^{-1} on the diagonal; $c_{ij} = (E_j/E_i)^{1/2}$ for site j in N_i, the set of neighbouring sites of site i, and 0 elsewhere, and $\phi \in (\phi_{\min}, \phi_{\max})$, where $\phi_{\min}$ and $\phi_{\max}$ are determined from the eigenvalues of C such that $M^{-1}(I - \phi C)$ is positive definite (Cressie, reference [13], p. 559). For this particular construction, $\phi_{\min}$ and $\phi_{\max}$ (and, indeed, all the eigenvalues of $I - \phi C$) are unaffected by a change in the set of expected counts $\{E_i\}$. To see this, let A represent a generic neighbour matrix with $a_{ij} = 1$ for site j in N_i, the set of neighbouring sites of site i, and 0 elsewhere. For any (non-zero) choice of $\{E_i\}$, $A = M^{-1/2}CM^{1/2}$, a similarity transformation; thus the set of eigenvalues of A equals the set of eigenvalues of C, from which the eigenvalues of $(I - \phi C)$ are uniquely determined. Stern and Cressie [15] show that the intrinsic autoregressive (IAR) models of Besag *et al.* [16] are obtained from (7) in the limit, as $\phi \uparrow \phi_{\max}$.

Modelling spatial structure involves a trade-off between the richness of H and Σ. We take the autoregressive coefficient matrix H to be parameterized by η_0, η_1 and η_2, such that

$$h_{ij} = \begin{cases} \eta_0 & \text{if } j = i \\ \eta_1 & \text{if } j \in N_i, \text{ that is, site } j \text{ is a neighbour of site } i \\ \eta_2 & \text{if } j \in N_i^{(2)}, \text{ that is, site } j \text{ is a second-order neighbour of site } i \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

Site j is a neighbour of site i if some proximity criterion is satisfied (for example, they share a common boundary), and we define site j to be a second-order neighbour of site i if $j \notin N_i$ but there exists a k such that $k \in N_i$ and $j \in N_k$. With this construction, we can interpret η_0 as a global measure of how much any site is affected by itself at one previous time lag, while η_1 and η_2 are global measures of the impact of first- and second-order neighbours, respectively, again at one previous time lag. The term $\{\boldsymbol{\varepsilon}_t\}$ in (5) captures instantaneous spatial correlation as well as the piecewise changes in mean given by (2).

We can assign prior distributions directly to the η_ℓ ($\ell = 0, 1, 2$), but it may be more sensible first to transform them as $\theta_\ell = f(\eta_\ell)$, and then to assign a prior to θ_ℓ. In the one-dimensional case, the time series (1) is stationary if $|\eta| < 1$. With this restriction, we see that $\theta = \log[(1+\eta)/(1-\eta)]$ is defined on $(-\infty, \infty)$, and a Gaussian prior on θ is reasonable if one expects unimodal variability in the autoregresive parameter. Otherwise, Gaussian mixture priors could be used. For the matrix H, we make an analogous transformation on its entries, choosing $\theta_\ell = \log[(1+\eta_\ell)/(1-\eta_\ell)]$, $\ell = 0, 1, 2$. We note, however, that this transformation from the η-scale to the θ-scale does not guarantee that the variances and covariances of the AR(1) process (5) are stationary. Since we are modelling a highly dynamic phenomenon through the non-constant function $\beta_{\rho(t)}$, stationarity is less of a concern. We could guarantee stationary variances and covariances by reparameterizing H through its eigenvalues, but that would destroy the natural interpretation present in our model, and we choose not to do so.

Using the model specified by equations (3)–(7), we can write the joint posterior of all parameters as proportional to the product of the likelihood and the priors. We note, however, that (5) is not valid for $t = 1$, since $\mathbf{s}_0$ is undefined, so we must explicitly define the distribution of $\mathbf{s}_1$. In most applications, the starting point $t = 1$ will be selected at a point where the disease stage is stable (that is, in stage 0, where there is no epidemic). If the process is stationary in this stage, the mean and variance of $\mathbf{s}_t$ are $(I - H)^{-1}\mathbf{1}\beta_0$ and $\sum_{k=0}^{\infty} H^k \Sigma (H')^k$, respectively. There is considerable computational overhead in calculating these quantities, so we avoid this problem and specify the distribution of $\mathbf{s}_1$ to be $\text{MVN}(\mathbf{0}, \zeta^2\Sigma)$, where $\zeta^2 > 1$ is chosen to reflect our additional uncertainty about $\mathbf{s}_1$. In the analysis of Section 3, we choose $\zeta^2 = 4$, and we note that the results are not particularly sensitive to this choice.

Using $[Y]$ as generic notation for the density of Y, the joint posterior distribution becomes

$$\begin{aligned} &[\boldsymbol{\alpha}, \beta_0, \beta_1, \beta_2, \sigma^2, \phi, \theta_0, \theta_1, \theta_2, \{s_{it}\} | \{y_{it}\}] \\ &\quad \propto \left(\prod_{t=1}^{T} \prod_{i=1}^{I} [y_{it} \,|\, \boldsymbol{\alpha}, s_{it}] \right) \times [\mathbf{s}_1 | \sigma^2, \phi] \\ &\qquad \times [\mathbf{s}_2 \,|\, H\mathbf{s}_1, \beta_{\rho(2)}, \sigma^2, \phi] \cdots [\mathbf{s}_T \,|\, H\mathbf{s}_{T-1}, \beta_{\rho(T)}, \sigma^2, \phi] \\ &\qquad \times [\boldsymbol{\alpha}][\beta_0][\beta_1][\beta_2][\sigma^2][\phi][\theta_0][\theta_1][\theta_2] \end{aligned} \tag{9}$$

Notice that our notation suppresses dependence of the posterior distribution on the explanatory variables $\{\mathbf{x}_{it}\}$, although it is understood that these make up part of the conditioning variables. We complete the Bayesian hierarchical model by specifying priors

$$\begin{aligned}
\boldsymbol{\alpha} &\sim \mathrm{MVN}(\mathbf{0}, \kappa^2 I) \\
\beta_\ell &\sim \mathrm{Gaussian}(\mu_\ell, \tau_\ell^2), \quad \ell = 0, 1, 2 \\
\theta_\ell &\sim \mathrm{Gaussian}(0, \delta_\ell^2), \quad \ell = 0, 1, 2 \\
\sigma^2 &\sim \text{Inverse gamma}(a, b) \\
\phi &\sim \mathrm{Uniform}(\phi_{\min}, \phi_{\max})
\end{aligned} \tag{10}$$

For a particular application, the parameter values of these prior distributions are specified; see Section 3 for the case of the influenza epidemic.

Fitting this model to any data set requires sophisticated numerical techniques. In the analysis of Section 3, we will employ Markov chain Monte Carlo. Cressie and Mugglin [10] lay out the necessary full conditional distributions for a similar model, and they also describe several computational issues related to its MCMC implementation.

3. AN INFLUENZA EPIDEMIC IN SCOTLAND AND ITS STATISTICAL ANALYSIS

During the winter of 1989–1990, Scotland experienced a large influenza epidemic. It is difficult to know how many people were infected, since most do not even visit their physicians, and records for those who do are generally not available. One estimate is that approximately 600 000 people (about one-tenth of the population of Scotland) were infected during this period, with perhaps only one in five visiting a physician (Christie P, 2000, personal communication). However, a fraction of the infecteds had severe enough complications that they had to be admitted to a hospital, and these cases are recorded and tabulated. During the winter of 1989–1990, there were approximately 500 emergency admissions where the primary cause was recorded as 'influenza' (ICD9 code 487). Henceforth, we shall focus on these emergency admissions, terming them 'cases,' while 'infecteds' will refer to those in the population who are ill with influenza but not necessarily hospitalized.

The smallest level of spatial aggregation available to us is the postcode sector. Each sector, of which there are 895 in Scotland, averages a population of about 5600 people. Temporally, we have available each case's date of admission. With only about 500 cases in all of Scotland in the winter of 1989–1990, this results in a data set with many zero counts. Thus we must aggregate in both space and time to some level where we can both deal with the sparseness of cases and yet not lose local spatial or temporal information. We chose to use as our spatial level of aggregation the 56 local government districts as used in the 1991 census; they are defined for the purposes of local government and not necessarily according to any natural or geographical features. The median population of a district is approximately 65 000, with a minimum of about 10 000 (in Nairn) and a maximum of about 650 000 (in Glasgow). For the neighbourhood structure $\{N_i\}$ in (8), we use the same neighbour definitions as those of Clayton and Kaldor [17] and given explicitly by, for example, Stern and Cressie [15]. We use weeks as our temporal level of aggregation. To align with covariates, as we shall discuss below, we allow the last week of December to have eight days, so that there are exactly 52

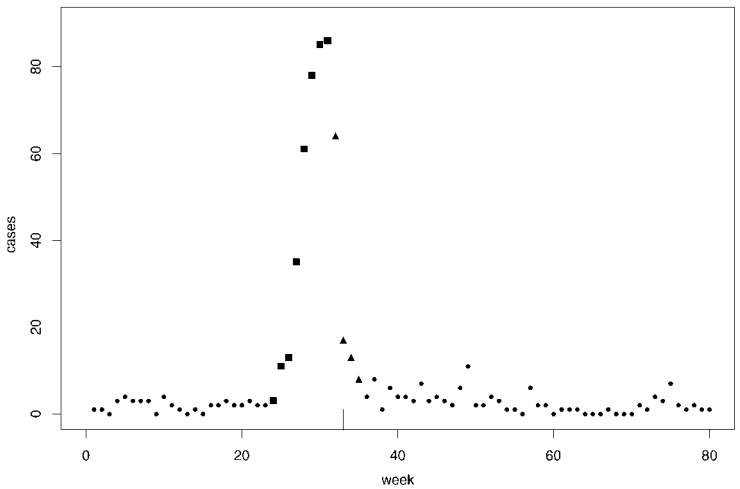

Figure 1. Number of cases (hospitalizations) in all of Scotland at each time point of the study period. Squares indicate a rising epidemic, while triangles indicate its decline. The vertical line segment on the horizontal axis at week 33 denotes the week of 1 January 1990; it is included as a reference point.

weeks in a year. The model can incorporate this through a slightly more complicated $\beta_{\rho(t)}$ in (2), though we view the extra generality as unnecessary in this setting.

For the purposes of this analysis, we focus on an 80-week period that occurs in the years 1989 and 1990. This period captures the largest epidemic seen in the ten-year period, 1986–1995. Somewhat arbitrarily, we select week 1 of our study period to be the 21st week of 1989 (corresponding to the week beginning 21 May 1989), so that the epidemic occurs between weeks 23 and 35, a period beginning the week of 22 October 1989 and continuing until the week of 15 January 1990. The last week of the study period ends on 2 December 1990, a choice that intentionally includes only one winter.

Figure 1 shows the total number of cases in all of Scotland at each time point (that is, during each week). We note again the many zeros in the data set, since when the epidemic is in stage 0 (that is, no epidemic), there are many weeks in which there is a total of 0 or 1 cases in all of the 56 districts of Scotland. Even at the peak during week 31, there are only 86 cases in the whole country.

The squares in Figure 1 indicate the weeks when the epidemic is rising, and we take these time points to be those in which the epidemic is in stage 1 (that is, a growing stage). The triangles represent the recession of the epidemic, and we take these points to belong to stage 2. In the notation of Section 2, we choose $t_0 = 24$, $t_1 = 32$ and $t_2 = 36$. At t_2, the stage returns to 0, and the epidemic recedes to vary about a mean that is stable. The dots indicate stage 0. The epidemic is potentially turned on at t_0 for all of Scotland, but of course not each district is affected immediately. How quickly the relative risk *appreciably* increases in any particular district depends on the matrix $(I - H)^{-1}\mathbf{1}$, which gives different behaviour for different districts. It also depends on the variability of $\mathbf{s}_t$ (since $\mathbf{s}_t$ is a random process), which depends on the matrix $(I - \phi C)^{-1}M$.

In calculating the expected number of cases E_i per week in each district, we adjust for demographic effects due to age and gender as follows. Our data set contains information on

gender as well as age (divided into 19 quinquennial age brackets: 0–4, 5–9, …, 85–89, 90+), yielding 38 demographic strata. We define $E_i = \sum_{k=1}^{38} R_{ik} q_k$, where R_{ik} is the population in district i and stratum k at risk of becoming a case, and q_k is the proportion in stratum k expected to become a case in any given week. To calculate E_i, we estimate q_k with

$$\widehat{q_k} = \frac{\sum_t \sum_{j=1}^{56} y_{jkt}}{\sum_t \sum_{j=1}^{56} R_{jkt}}; \quad k = 1, \ldots, 38$$

where y_{jkt} is the number of observed cases in district j in the kth stratum during week t. To obtain y_{jkt}, we observe cases in Scotland during the years 1986–1995, much longer than the study period in question, so that the summation over t in the formula above covers $52 \times 10 = 520$ weeks. In these years, we assume that the population at risk is constant and equal to the value recorded in the 1991 census; hence R_{jkt} does not in fact depend on t. The resulting age- and sex-adjusted weekly expected cases per district $\{E_i\}$ are quite small, with a median value of 0.066, a minimum of 0.0104 (in Badenoch/Strathspey), and a maximum of 0.63 (in Glasgow).

Influenza outbreaks almost always occur in late autumn and winter. The most common explanations given are that in colder temperatures the virus is able to live longer and that the virus spreads more easily because people tend to stay indoors and are in closer contact with each other. Temperature is not the *cause* of an epidemic, but it may be a contributing factor, so we include it as a covariate in our model. Temperature data are only available by weekly averages from the cities of Edinburgh, Glasgow and Aberdeen. We imputed these temperature data to each of the remaining districts using inverse-distance-squared weighted averages (for example, Cressie, reference [13], p. 371), where distances are computed between district centroids. Exploratory data analysis does not reveal any strong suggestions as to *when* temperature has its effect (if any); intuition suggests that, given the incubation period of influenza, the number of infecteds might rise about a week after a drop in temperature, and a rise in the number of cases (emergency admissions) might lag the rise in infecteds by several more days. To capture this intuition, we include temperature at time lags of 0, 1 and 2 weeks as covariates in our model.

In addition, we account for districts that seem to be sicker in general than others. We have available to us the *all-cause* standardized mortality ratio for the years 1989 and 1990, by district. We average the two years' values by district and include them as a spatially varying covariate.

With these covariates, we now expand equation (4) to read

$$z_{it} = \mathbf{x}_{it}'\boldsymbol{\alpha} + s_{it} = \alpha_0 \Upsilon_{i,t} + \alpha_1 \Upsilon_{i,t-1} + \alpha_2 \Upsilon_{i,t-2} + \alpha_3 \,\mathrm{SMR}_i + s_{it} \tag{11}$$

where $\Upsilon_{i,t}$ denotes the temperature in district i at week t, and SMR_i denotes the all-cause standardized mortality ratio in district i.

3.1. Fitting the spatio–temporal hierarchical model

We now discuss the choice of priors in (10) and the practical difficulties encountered in obtaining the posterior distributions. Based on both exploratory data analysis and our desire to choose vague priors, we select values for the hyperparameters as in Table I. To obtain the posteriors, there are a number of technical challenges related to the implementation of the

Table I. Table of selected parameters, showing prior distributions and prior and posterior summaries.

Parameter	Distribution	Prior quantiles			Posterior quantiles		
		0.025	0.5	0.975	0.025	0.5	0.975
α_0	Gaussian(0,4)	−3.92	0.0	3.92	−0.0692	−0.0032	0.0628
α_1	Gaussian(0,4)	−3.92	0.0	3.92	−0.0864	−0.0143	0.0518
α_2	Gaussian(0,4)	−3.92	0.0	3.92	−0.1185	−0.0508	0.0193
α_3	Gaussian(0,4)	−3.92	0.0	3.92	−0.0075	−0.0018	0.0031
β_0	Gaussian(0,4)	−3.92	0.0	3.92	−0.9945	−0.2618	0.0204
β_1	Gaussian(0,4)	−3.92	0.0	3.92	0.7073	1.4505	2.9803
β_2	Gaussian(0,4)	−3.92	0.0	3.92	−0.3022	0.6874	3.2501
θ_0	Gaussian(0,4)	−3.92	0.0	3.92	0.0988	0.5122	0.9271
θ_1	Gaussian(0,4)	−3.92	0.0	3.92	−0.7222	0.1008	0.4575
θ_2	Gaussian(0,4)	−3.92	0.0	3.92	−0.1043	0.0045	0.0983
σ^2	Inverse gamma(0.25,0.4)	1.457	57.2	9.5×10^6	0.0579	0.081	0.1065
ϕ	Uniform$(-0.325, 0.1752)$	−0.3125	−0.075	0.1627	0.1022	0.1514	0.1663

MCMC algorithm. We note the large number of parameters (there are 4480 $\{s_{it}\}$ variables) and, due to high autocorrelations, the need to run the sampler for many iterations. Running five parallel independent chains for approximately 10 000 iterations each required more memory than we had available in our workstation, so we had to access the disk as virtual memory, a process that further (and dramatically!) slowed down the computation. Implemented in the C programming language, this algorithm required about 20 hours to run on a Sun Ultra Enterprise 250 workstation. Convergence was monitored via the Gelman and Rubin [18] convergence diagnostic and visual inspection of trace plots. Convergence was obtained within the first 2000 iterations, and these were then discarded, leaving approximately 8000 iterations of five chains each, for a total of just over 40 000 observations from the posterior distribution of each parameter. Of course, the autocorrelation in the Markov chains leads to a smaller effective sample size; see, for example, the comment by Neal in Kass *et al.* [19]. For the 4492 parameters considered, the effective sample sizes ranged from 200 to 7000, leaving us confident that our MCMC inferences are valid. Cressie and Mugglin [10] gave an independent check of this when they recovered, from MCMC inference, known parameter values used in their simulation.

Posterior medians and 95 per cent Bayes credible intervals are shown in Table I. We note in particular that the posteriors for all parameters are much tighter than the priors, so despite the many zeros in the data set there is substantial information to infer about these parameters. The posterior medians of β_0, β_1 and β_2 are -0.26, 1.45 and 0.69, respectively, agreeing with our intuition that the highest value for the mean of the epidemic-forcing term will occur when the epidemic is in stage 1, the second highest will be in stage 2, and the lowest will be in stage 0. Furthermore, we note that the only appreciable (in the sense that 0 is not in the 95 per cent posterior credible interval) autoregression parameter in (8) is $\theta_0 = \log[(1+\eta_0)/(1-\eta_0)]$, the term corresponding to lag-1 influence of a district on itself, and it is positive. Of the regression parameters $\{\alpha_\ell\}$ in (11), none is appreciable, though α_2 suggests a negative value, as seen in both Table I and Figure 2. This indicates that a drop in temperature at a two-week

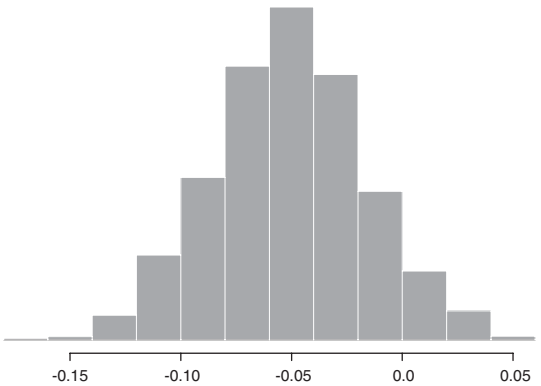

Figure 2. Histogram of posterior samples for regression parameter α_2, suggesting a negative value for the parameter.

lag correlates somewhat with an increased relative risk of cases, agreeing with our intuition given earlier in this section that cases may lag infecteds, which may in turn lag a temperature drop. Figure 2 also exemplifies the well known advantage of the Bayesian approach in that we obtain the entire (sampled) posterior distribution of all parameters and not just a point or interval estimate.

To assess the performance of our method, we apply a goodness-of-fit measure as proposed in Cressie and Mugglin [10], who simulate from the model they are attempting to fit. Thus they know the true values of the parameters they are attempting to estimate, and they consider the quantity

$$G_i \equiv \sum_t |\, s_{it}^* - \tilde{s}_{it}|/\sum_t 1$$

where s_{it}^* is the true value, and $\tilde{s}_{it}$ is the median of (the samples from) the posterior distribution of s_{it}. A measure of uncertainty that does not depend on knowing the true values $\{s_{it}^*\}$ is given by

$$W_i \equiv \sum_t w_{it}/\sum_t 1$$

where w_{it} is the width of the 95 per cent Bayesian credible interval from the posterior distribution of s_{it}, a quantity that is roughly proportional to the posterior standard deviation. The authors demonstrated a strong linear relationship between $\{G_i\}$ and $\{W_i\}$, so that W_i serves as a proxy for goodness-of-fit as well as a measure of variability.

Intuitively, when there are many cases in a district, the available information is high, and the posterior distributions of parameters for these districts will exhibit small variances. When there are few cases, the opposite will occur. Thus, in a model that fits well, we expect an inverse relationship between E_i and G_i. Moreover, we expect the variance of the ratio of observed to expected cases in district i to be proportional to $1/E_i$, and we further expect the posterior variance of s_{it} to be proportional to the same quantity. Thus, with an expected linear relation between $\{G_i\}$ and $\{W_i\}$, we expect a linear relationship between W_i and the quantity $A_i \equiv 1/\sqrt{E_i}$. Departures from linearity may indicate districts that do not behave according to some aspect of the model.

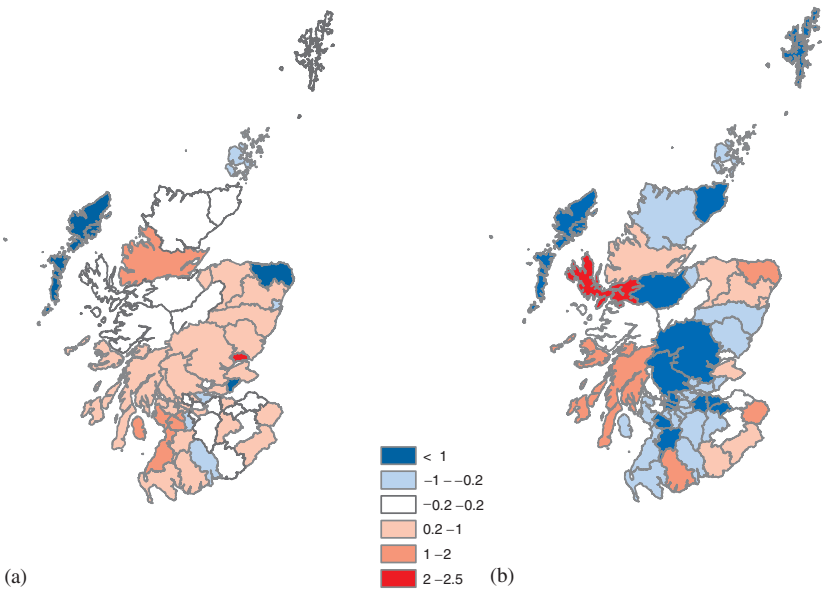

Plate 1. Comparison of velocity of log relative risk change at weeks 28 (left map) and 32 (right map). Red indicates a positive value, while blue indicates a negative value.

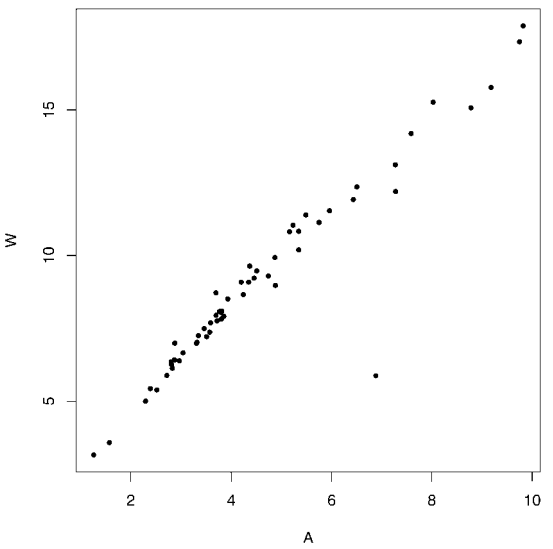

Figure 3. Goodness-of-fit measure W_i plotted against $A_i \equiv 1/\sqrt{E_i}$.

Figure 3 plots W_i against A_i. The leftmost two points (where there are many cases) come from the cities of Glasgow and Edinburgh, while the rightmost two (where there are few cases and few expected cases) come from Nairn and Badenoch/Strathspey. We note the clear linearity in the plot, indicating a good overall fit. The outlier represents the Shetland Islands, which experienced an inordinately high number of cases (7 cases). Four of the cases appeared in the epidemic window, but the other 3 cases appeared in weeks 1, 50 and 54; there were no other cases in the rest of Scotland in weeks 1 and 54, and only one other case in week 50. Furthermore, since the Shetland Islands only had one neighbour, there is substantial spatial smoothing resulting from relatively small prior variability. The end result is a W-value for the Shetland Islands that is unusually small.

3.2. Interpretation of derived quantities

As was pointed out in Section 1, in the infectious-diseases setting, we are very interested in the spatio–temporal dynamics and their implications. With samples from the posterior distributions of all parameters, we can examine almost any derived quantity of interest.

We illustrate first a well known feature of Bayesian spatial modelling, namely that the posterior relative risks represent smoothed values of the raw standardized morbidity ratios (SMRs) $\{y_{it}/E_i\}$. Consider Figure 4, which indicates in Figure 4(a) the raw SMRs and in Figure 4(b) the relative risks across Scotland for week 27. Week 27 is in the middle of the growth of the epidemic (stage 1), and yet there are only 35 cases in all of Scotland during this week; 36 of the 56 districts report no cases, giving them an observed SMR of zero. This is reflected in the white regions of Figure 4(a). Twelve districts report one case each, while the other eight report the remaining 23 cases. Despite the many zero counts, there is obviously an epidemic in progress, and no one should really believe that there is no risk of disease in these zero-count districts. It is much more sensible to imagine an underlying smooth relative

Figure 4. (a) Standardized morbidity ratio, $y_{i,27}/E_i$, during week 27. (b) Relative risk $e^{z_{i,27}}$ during week 27.

risk, from which the observed cases represent one possible realization of what might occur. Such a relative risk structure is indicated in Figure 4(b), where we map the posterior medians of the relative risk for week 27. We note that the map is generally smoother, with obvious spatial similarity between neighbouring districts, and yet the overall essence of the spatial distribution of the SMRs $\{y_{i,27}/E_i\}$ is well captured.

Our Bayesian model is dynamically spatial, so we can examine posterior quantities that represent the temporal progression of the disease. We illustrate this by examining a series of histograms of log relative risks for one particular district (Glasgow, which in our enumeration corresponds to $i = 41$). Glasgow is the largest city in Scotland, and its expected count, E_{41}, is the largest. Figure 5 shows a series of histograms of log relative risk during various weeks of the epidemic. We observe the log relative risk $\{z_{41,t}\colon\ t = 26, \ldots, 33\}$ growing from a value near zero up to the range of two to four, then receding again as the epidemic abates. In addition, we note that the widths of the histograms are smallest during the height of the epidemic. This reflects the higher level of certainty derived from observing more cases, and the lower level of certainty when the number of cases is small.

While Figure 4 displays spatial smoothing and Figure 5 displays dynamics, we can also analyse and display the spatial and temporal progression of the underlying relative risks. To do this, we introduce 'bubble' maps in Figure 6. In each map, we draw a circle centred at the centroid of each district, where the area represents the value of the relative risk in the district. To mitigate perception bias, we choose the areas of the circles to be proportional to the relative risk raised to the power of 3/2; Stevens' Law (Stevens [20]) states roughly

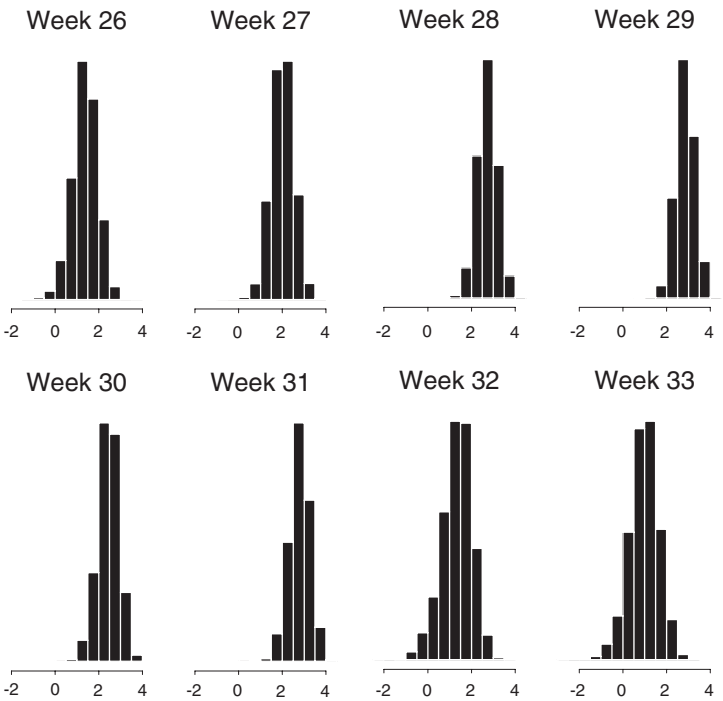

Figure 5. Histograms of log relative risk for Glasgow during weeks 26–33 of the epidemic.

Figure 6. 'Bubble' map of relative risks (posterior medians) during the height of the epidemic. Circles are centred at district centroids, with areas indicating value of relative risk.

that people perceive the relative areas of shapes as the ratio of the areas raised to a power approximately equal to 2/3.

While perhaps not as aesthetically pleasing as choropleth maps, these maps have several advantages: they remove extraneous information, such as boundaries; they avoid the misperception that a large district area implies a large value or importance; they can display a continuous variable via circle areas, rather than being restricted to a few colour ranges; they are easy to generate and do not consume as much computer disk space.

Figure 6 shows the progression of the disease in both time and space. In week 26, there is some mild elevated relative risk visible in a few districts, notably one south of Glasgow (Cumnock and Doon Valley) and another in Dundee. In week 27, the disease appears to be present mostly in a southwest–northeast direction, in a corrider of high population density that includes the cities of Glasgow, Edinburgh, Dundee and Aberdeen. Week 28 exhibits the same general spatial trend, though there has been some growth in a few districts. By week 29, the disease appears to be spreading outward, approximately perpendicular to the above-mentioned corridor, while in the corridor itself there is some attenuation. By week 30, the growth in the outer regions is pronounced, and at week 31, several districts that peaked in week 30 are beginning to decrease, while others appear to be peaking. At week 32, the disease is mostly along the perimeter of Scotland, and generally declining, especially so in the geographical centre. Week 33 exhibits marked decline, with the remaining risk distributed along the eastern coastline.

The dynamics revealed in Figure 6 lead us to ponder the question of how fast the disease is spreading in space and growing in time. To assess temporal velocity, we examine a difference of log relative risks. Consider the log relative risk vector $\mathbf{z}_t$ at time t, and define the velocity $\mathbf{v}_t$ as

$$\mathbf{v}_t \equiv \mathbf{z}_t - \mathbf{z}_{t-1} = (\mathbf{x}_t - \mathbf{x}_{t-1})'\boldsymbol{\alpha} + (\mathbf{s}_t - \mathbf{s}_{t-1}) \tag{12}$$

We stress that this is a measure of change in log-relative risk in one time period, rather than a measure of spread per unit distance. Using the existing MCMC simulations, we can compute posterior samples from the distributions of the $\mathbf{v}_t$'s; the resulting posterior medians are mapped in Plate 1. The figure indicates maps of velocity in weeks 28 (when the epidemic is growing) and 32 (when the epidemic is receding). Note from Figure 1 that the overall disease prevalence is similar in these two weeks. In Plate 1, a red shade indicates a positive velocity (that is, growth of log relative risk), while a blue shade indicates negative velocity. Thus, in week 28 there is generally positive velocity (red tints) in the higher population corridor, with a few blue and white regions on the periphery. By week 32, when the epidemic is generally receding, we observe some districts with markedly decreasing velocities, while others (especially on the periphery) are still increasing. This agrees well with our observations regarding Figure 6.

We now consider the impact of the epidemic-forcing term $\boldsymbol{\varepsilon}_t$ and its mean $\beta_{\rho(t)}\mathbf{1}$ when the epidemic undergoes a stage transition. Given a β-change that uniformly affects all of Scotland (as perhaps induced by climatic conditions or movement of many people during a holiday period), what effect does a change $\Delta\beta$ have on the dynamic spatial process? We can quantify this by computing the eventual change $\Delta\boldsymbol{\mu}$ in the mean of $\{\mathbf{s}_t\}$ caused by $\Delta\beta$. Recall from (5) that $\mathbf{s}_t = H\mathbf{s}_{t-1} + \boldsymbol{\varepsilon}_t$. A stationary mean $\boldsymbol{\mu}$ would satisfy $\boldsymbol{\mu} = H\boldsymbol{\mu} + \beta\mathbf{1}$, where $E(\boldsymbol{\varepsilon}_t) = \beta\mathbf{1}$. Solving for $\boldsymbol{\mu}$, we obtain

$$\boldsymbol{\mu} = (I - H)^{-1}\mathbf{1}\beta \tag{13}$$

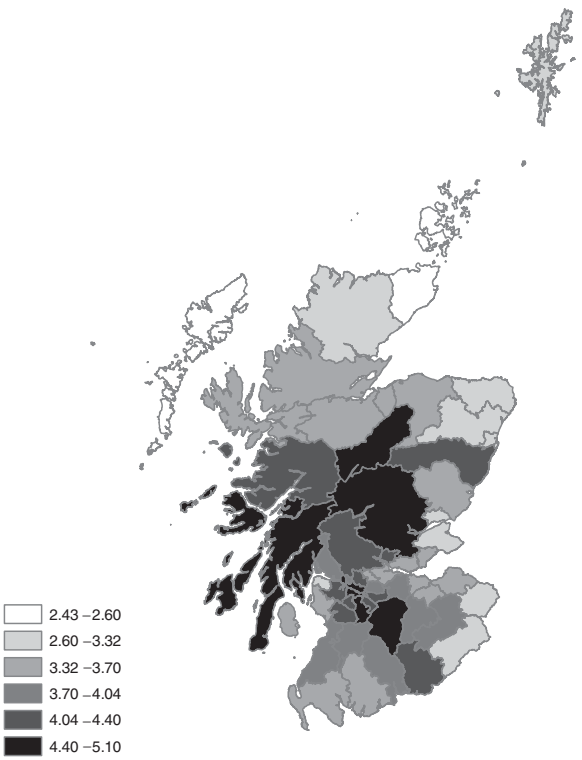

Figure 7. Eventual change $\Delta\boldsymbol{\mu}$ caused by a change $\Delta\beta = \beta_1 - \beta_0$ in the mean forcing term.

and hence

$$\Delta\boldsymbol{\mu} = (I - H)^{-1}\mathbf{1}\Delta\beta \tag{14}$$

Note that $(I - H)^{-1}$ is not guaranteed to exist for every realization from the posterior of (η_0, η_1, η_2); in our case it did, though with a few realizations for which at least one eigenvalue of $I - H$ was close to 0, resulting in a posterior distribution of $\Delta\boldsymbol{\mu}$ with rather heavy tails.

Figure 7 shows a map of the posterior medians of $\Delta\boldsymbol{\mu}$ evaluated for $\Delta\beta = \beta_1 - \beta_0$, which occurs at $t = t_0 = 24$. It is striking that a uniform β-change over Scotland results in such a differential impact. We observe the spatial structure, with the regions of highest impact occurring in two clusters, one each to the north and south of the high-population corridor. These are the regions that, when the conditions suddenly turn to favour an epidemic, are the most profoundly affected; if allowed to stay in stage 1, they would experience the biggest rise in disease prevalence.

3.3. Assessing what might have happened

Since we view the data as one possible realization from a hidden process, we wonder then what other scenarios might have occurred. Was this epidemic unusually bad? Might it have

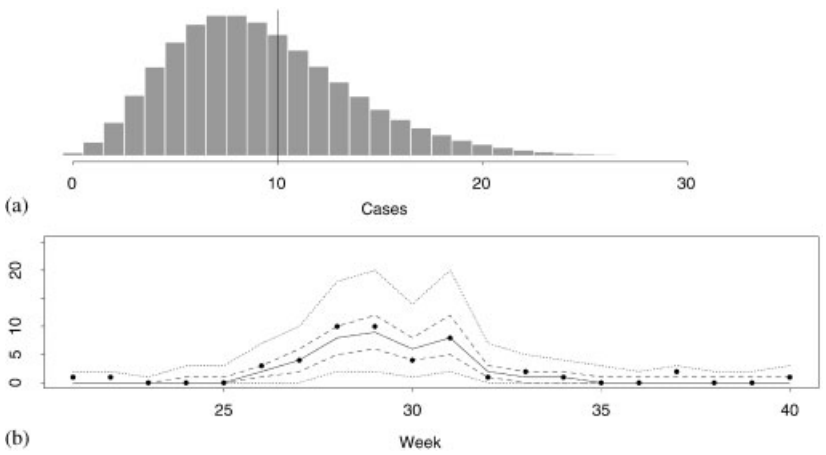

Figure 8. (*a*) Histogram of posterior predictive distribution for Glasgow at week 29. The vertical line represents the observed data value. (*b*) Posterior predictive distribution for $y^* \mid y$ over weeks 21 to 40 of the epidemic. Lines indicate the 0.025, 0.25, 0.5, 0.75 and 0.975 quantiles, while dots indicate the data values (y).

been much worse? If so, what planning might have been wise for hospitals or other health facilities?

A mathematical framework for exploring this question is the notion of a *posterior predictive distribution* (for example, Rubin [21] and Gelman *et al.* [22]), in which we calculate the distribution of (hypothetical) cases given the observed data, $[\{y^*_{it}\} \mid \{y_{it}\}]$, integrating out all other parameters. In a similar notation to that given in Section 2, we suppress dependence of the posterior predictive distribution on the explanatory variables $\{\mathbf{x}_{it}\}$, although it is understood that these make up part of the conditioning variables. Notice that this distribution is 'predictive' only in the sense of predicting other possible realizations from the spatio–temporal hierarchical model; it does not predict how the epidemic is likely to evolve. An analytical derivation is difficult, but we can obtain samples from this distribution by generating observations from $[\{y^*_{it}\} \mid \{y_{it}\}, \boldsymbol{\theta}]$, where $\boldsymbol{\theta}$ now represents *all* unknowns, and where we choose as our values for $\boldsymbol{\theta}$ samples from the full posterior distribution $[\boldsymbol{\theta} \mid \{y_{it}\}]$ given by (9). This we did for Glasgow as a representative district, generating Poisson observations for the approximately 40000 posterior observations of $[\boldsymbol{\theta} \mid \{y_{it}\}]$, and we summarize this sample with histograms and sample quantiles in Figure 8.

Figure 8(*a*) shows this posterior predictive distribution of cases for Glasgow at week 29, where the vertical line represents the ten observed cases. Figure 8(*b*) shows the dynamic aspect of this process over the weeks 21 to 40, where the dots indicate the observed number of cases and the lines indicate the median, quartiles and 2.5 and 97.5 percentiles of the posterior predictive distributions. This indicates, first of all, that the model fits quite well, at least for Glasgow. It also indicates, on a weekly basis, what *might* have happened. For instance, in weeks 29 or 31, there could easily have been about 20 cases, or fewer than 5. The need for ten more hospital beds may be a minor problem, but considering the number of infecteds in the population, planning for other health services might be improved with a

knowledge about variability and worst-case scenarios. In particular, there is some evidence that, during an influenza epidemic, emergency admissions for causes such as other respiratory conditions and heart disease also increase substantially (Simonsen *et al.* [23]). This type of analysis can easily be used for any district, and it can be extended to cover any aggregation of districts or weeks, up to and including all cases in all of Scotland in the entire 80-week period.

4. DISCUSSION

The Bayesian hierarchical model developed in this application attempts to capture both instantaneous spatial dependence as well as diffusion and growth in space and time by modelling the (log) relative risks in terms of spatial proximities. Our model has its limitations, as we have mentioned earlier. The most important one is that caused by both spatial and temporal aggregation, which prevents us from seeing how infected individuals interact with other individuals in their communities to spread the disease. Thus, we cannot claim to have captured the aetiology of the influenza epidemic in Scotland in 1989–1990, but rather we have a smoothed and interpretable description of what happened during the epidemic. One way to modify the current approach would be to start with a simple diffusion equation of disease spread, such as found in Diekmann and Heesterbeek (reference [24], p. 126), and look at its properties under discretization and aggregation. If one were successful in establishing a hierarchical model like that of (3)–(6), then the interpretable diffusion parameters would be estimable from the aggregated data; Lawson and Leimich [9] make an analogous point for their descriptive, individual-level, space–time models. We believe that the diffusion approach will not be easy, although it is certainly worthy of further research.

One might also ask how the results from a model of data that have been aggregated differently would compare to those given in Section 3 (for example, postcode sectors/months versus districts/weeks). Unless the aggregated model is built from a point-level diffusion model, it is very difficult to say how the epidemic would behave at higher or lower levels of aggregation. This problem is ubiquitous in small-area estimation and is known as ecological bias in the epidemiology community (for example, Greenland and Robins [25]) and the modifiable areal unit problem in the geography community (for example, Fotheringham and Wong [26]).

The analysis in Section 3 depends on certain modelling choices we made. In particular, the choice of neighbourhood criteria that define first- and second-order neighbour sets will affect both instantaneous correlation and spatial spread of the epidemic. We used the neighbourhood structure of Clayton and Kaldor [17], which is based on sharing district boundaries. (A listing of the 56 neighbourhoods can be found in Stern and Cressie [15].) For an infectious disease it might also be informative to use a neighbourhood structure that defines neighbours based on population movement. For instance, it may be that residents of the Shetland Islands rarely go to the Orkney Islands (their closest neighbours), but rather they may travel more frequently by sea or air to Aberdeen. If this is the case, a neighbourhood structure that accounts for effects of this type would be sensible.

There are two principal theories regarding what causes influenza outbreaks in Scotland. One theory suggests that the virus enters the country from Norway or England as part of a worldwide epidemic wave. Another posits that the virus is present all year in a dormant, possibly mutating form, and some unknown trigger sets it off. In either case, some unobserved

mechanism determines whether and when the epidemic will occur. For our analysis, we chose the change points t_0, t_1 and t_2 based on exploratory analysis, as in Figure 1. The duration of an outbreak is typically 10–12 weeks, regardless of the height of the peak (Christie P, 2000, personal communication), though some epidemics appear somewhat more symmetrical than this one. The object of the analysis in this paper is to smooth out the Poisson noise and look at the evolution of the epidemic through the (log) relative risk. Our approach is Bayesian so that the prior model that specifies global behaviour on regime changes is always modified by the local behaviour determined by the data. We reiterate that our analysis is descriptive, taking into account the whole of the data set to calculate the posterior distributions, rather than using only past and present information at any given time point to predict the epidemic's future evolution. To be able to forecast, we would need to build a stochastic model for t_0 (and possibly for other change points $t_1, t_2, \ldots$). In a climatological study of sea-surface temperatures, where some physics is known, Berliner *et al.* [27] build such a stochastic model for regime changes and use it in a hierarchical Bayesian statistical model that yields seven-month forecasts.

With the natural extension of spatial models to spatio–temporal models comes the need to produce maps that evolve in time. While display of this added dimension remains a presentational challenge, the use of a geographic information system (GIS) can greatly facilitate the quality and consistency of drawing maps. In addition, calculations regarding district centroids, areas, perimeters and contiguous boundaries are easily made within a GIS. We used some of these functionalities in the current study. Practitioners of spatial statistics are becoming increasingly acquainted with the use of GIS, though the statistical capability of a GIS is still generally limited to descriptive summaries rather than inference. Nevertheless, our Bayesian hierarchical statistical analysis has produced optimal estimates, and the GIS has proved to be a powerful tool for presenting our results.

ACKNOWLEDGEMENTS

The authors wish to thank Dr Peter Christie, Consultant Epidemiologist, Scottish Centre for Infection and Environmental Health, Glasgow, for many helpful insights into influenza epidemics. We would also like to thank the referees for their constructive comments on an earlier version of the manuscript. This research was supported by the U.S. Environmental Protection Agency under Assistance Agreement R827-257-01-0, and by the Office of Naval Research under grant number N00014-99-1-0001.

REFERENCES

1. Anderson RM, May R. *Infectious Diseases of Humans*: *Dynamics and Control*. Oxford University Press: London, 1992.
2. Jacquez JA. *Compartmental Analysis in Biology and Medicine*. 3rd edn. Ann Arbor BioMedware: Ann Arbor, 1996.
3. Bolker BM, Grenfell B. Impact of vaccination on the spatial correlation and persistence of measles dynamics. *Proceedings of the National Academy of Sciences* 1996; **93**:12648–12653.
4. Keeling MJ, Rand DA, Morris, AJ. Correlation models for childhood epidemics. *Proceedings of the Royal Society of London* 1997; **264**:1149–1156.
5. O'Brien SJ, Christie P. Do CuSums have a role in routine communicable disease surveillance? *Public Health* 1997; **111**:255–258.
6. Waller LA, Carlin BP, Xia H, Gelfand, AE. Hierarchical spatio-temporal mapping of disease rates. *Journal of the American Statistical Association* 1997; **92**:607–617.
7. Knorr-Held L, Besag JE. Modelling risk from a disease in time and space. *Statistics in Medicine* 1998; **17**: 2045–2060.
8. Cressie N, Stern HS, Wright DR. Mapping rates associated with polygons. *Journal of Geographical Systems* 2000; **2**:61–69.

9. Lawson AB, Leimich P. Approaches to space-time modelling of infectious disease behaviour. *IMA Journal of Mathematics Applied to Medicine and Biology* 2000; **17**:1–13.
10. Cressie N, Mugglin AS. Spatio-temporal hierarchical modelling of an infectious disease from (simulated) count data. In *COMPSTAT*: *Proceedings in Computational Statistics*, Bethlehem JG, van der Heijden PGM (eds). Physica-Verlag: Heidelberg, 2000; 41–52.
11. Carrat F, Valleron A-J. Epidemiologic mapping using the 'kriging' method: applications to an influenza-like illness epidemic in France. *American Journal of Epidemiology* 1992; **135**:1293–1300.
12. Gilks WR, Richardson S, Spiegelhalter DJ (eds). *Markov Chain Monte Carlo in Practice*. Chapman and Hall: London, 1996.
13. Cressie N. *Statistics for Spatial Data*. Revised edn. Wiley: New York, 1993.
14. Cressie N, Chan NH. Spatial modelling of regional variables. *Journal of the American Statistical Association* 1989; **84**:393–401.
15. Stern HS, Cressie N. Inference for extremes in disease mapping. In *Disease Mapping and Risk Assessment for Public Health*, Lawson A, Biggeri A, Bohning D, Lesaffre E, Viel J-F, Bertollini R (eds). Wiley: Chichester, 1999; 63–84.
16. Besag J, York JC, Mollié A. Bayesian image restoration, with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics* 1991; **43**:1–59.
17. Clayton D, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* 1987; **43**:671–681.
18. Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. *Statistical Science* 1992; **7**:457–511.
19. Kass RE, Carlin BP, Gelman A, Neal RM. Markov chain Monte Carlo in practice: a roundtable discussion. *American Statistician* 1996; **52**:93–100.
20. Stevens SS. *Psychophysics*. Wiley: New York, 1975.
21. Rubin DB. Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Annals of Statistics* 1984, **12**:1151–1172.
22. Gelman A, Meng X-L, Stern HS. Posterior predictive assessment of model fitness via realized discrepancies (with discussion). *Statistica Sinica* 1996; **6**:733–807.
23. Simonsen L, Clarke MJ, Williamson G, Stroup DF, Arden NH, Schonberger LB. The impact of influenza epidemics on mortality: introducing a severity index. *American Journal of Public Health* 1997; **87**:1944–1950.
24. Diekmann O, Heesterbeek JAP. *Mathematical Epidemiology of Infectious Diseases*: *Model Building*, *Analysis*, *and Interpretation.* Wiley: New York, 2000.
25. Greenland S, Robins J. Ecological studies: biases, misconceptions, and counterexamples. *American Journal of Epidemiology* 1994; **139**:747–760.
26. Fotheringham AS, Wong DWS. The modifiable areal unit problem in multivariate statistical analysis. *Environment and Planning A* 1991; **23**:1025–1044.
27. Berliner LM, Wikle CK, Cressie N. Long-lead prediction of Pacific SSTs via Bayesian dynamic modelling. *Journal of Climate* 2000; **13**:3953–3968.

