Bayesian approaches to epidemiological surveillance: a review and introduction for risk-assessors and decision-makers.

Working example 3:

A full Bayes approach to spatio-temporal surveillance for bluetongue in Norwegian cattle.

Our response variables are the number of late abortions a_{ij} and the number of deaths among cattle d_{ij} at week *i* in region *j.* We model these as Poisson variables:

$$a_{ij} \sim Po\left(b_{ij} + \delta_{(i-L_a)j} s_{(i-L_a)j}\right)$$

$$d_{ij} \sim Po\left(c_{ij} + \delta_{(i-L_d)j} k_{(i-L_d)j}\right)$$

where the baseline rates b_{ij} and c_{ij} , the disease indicators δ_{ij} (with possible values 0 or 1), and the additional rates s_{ij} and k_{ij} are described below. Note that we allow for a time delay between the infection and the resulting syndromes. Here we have used L_a = 3 and L_d = 1. The baseline rates are modelled with

$$\log\left(b_{ij}\right) = \log(\alpha_a) + \log\left(x_{(i-L_b)j}\right)$$

$$\log\left(c_{ij}\right) = \log(\alpha_d) + \log\left(y_{ij}\right) + \varepsilon_{ij}$$

where x_{ij} is the number of cattle days for cattle in late pregnancy and y_{ij} is the number of living days for cattle. Note that the baseline rate of late abortions is modelled as proportional to the number of cattle days for cattle at late pregnancy L_b weeks earlier. Based on initial data analysis, we used the estimate $L_b = 7$. The baseline rate of deaths is modelled as proportional to the number of living days for cattle with random effects $\varepsilon_{ij} \sim N(0, \sigma^2)$. However, as a simplification in this example, we assume the random effects are independent. As a further simplification, we have omitted the random effect for the late abortion baseline.

Increase in late abortions and increase in mortality when $\delta_{ij} = 1$ (grid infected) are modelled as

$$\log\left(s_{ij}\right) = \log(\theta_a) + \log\left(x_{ij}\right)$$

$$\log\left(k_{ij}\right) = \log(\theta_d) + \log\left(y_{ij}\right)$$

Again, we have for simplicity omitted random effects.

For syndromic surveillance the infection status δ_{ij} is unknown. We assume several independent possibilities for infection: Each week, there is a probability π_1 for a new outbreak, a probability π_2 for an infection spreading from each previously infected neighbouring region, and a probability π_3 for an infection spreading from each previously infected non-neighbouring region. Specifically,

$$P\left(\delta_{i+1,j} = 0 \middle| \delta_{ij} = 0, z_{ij}, w_{ij}\right) = (1-\pi_1)(1-\pi_2)^{z_{ij}}(1-\pi_3)^{w_{ij}}$$

where z_{ij} and w_{ij} are the number of infected neighbours and non-neighbours, respectively. We assume an absorbing state model, so that

$$P\left(\delta_{i+1,j} = 1 \middle| \delta_{ij} = 1, z_{ij}, w_{ij}\right) = 1$$

It is assumed that no grid is infected before the syndromic surveillance so the initial probability of infection is π_1 for each grid point.

The prior distributions used are listed below.

$$\log(\alpha_a) \sim N(0, 100^2)$$

$$\log(\alpha_d) \sim N(0, 100^2)$$

$$\log(\theta_a) \sim N(0, 100^2)$$

$$\log(\theta_d) \sim N(0, 100^2)$$

$$\sigma \sim U(0{,}100)$$

$$\pi_1 \sim Beta(1{,}2000)$$

$$\pi_2 \sim Beta(1{,}20)$$

$$\pi_3 \sim Beta(1{,}40)$$