Integrating wildlife movement models with infectious disease dynamics

Chris Jewell, Kyle Richardson, Dan Tompkins, Bryce Buddle, Nigel French

Massey University
c.p.jewell@massey.ac.nz

DEPE 2013, Wellington

Overview

Motivation

- Bovine tuberculosis
  - Cost
  - Maintenance host: brushtail possum

- Understand dynamics of bTB transmission in possum populations
  - How fast?
  - Determinants (ie. which possums)?

- Can we use trapping data, proximity logging, lab testing to infer disease dynamics?

Approach

- Data
  - Capture-recapture data (movement)
  - Proximity loggers (contact network)
  - Lab sampling (infection status)

- Bayesian epidemic models
  - Stochastic dynamical models
  - Flexible – may include many sources of data
  - Reflect underlying population characteristics
  - Account for missing data

Epidemic models: history
The SIR Model

Susceptible $\rightarrow$ Infected $\rightarrow$ Removed

$$
\begin{aligned}
\frac{dS}{dt} &= -\beta SI \\
\frac{dI}{dt} &= \beta SI - \gamma I \\
\frac{dR}{dt} &= \gamma I
\end{aligned}
$$

**Refs**: Kermack and McKendrick (1927); Bailey (1975); Becker (1989)

Population-based approach

Models transmission between pairs of individuals

1. **Flexibility** of modelling
   - Detailed heterogeneity
   - Interpretability of parameters

2. Incorporation of **different types of data**
   - Probabilities
   - Rates
   - Contact matrices

Allows modelling of highly heterogeneous populations at the **individual** (farm) level
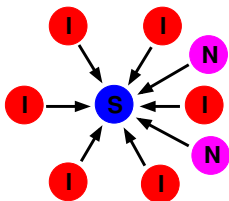
The Model
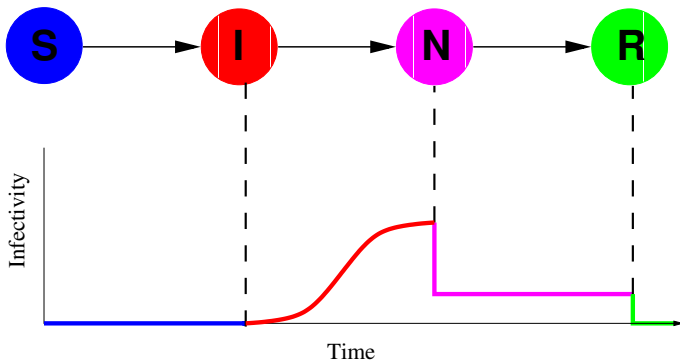
- Continuous time **stochastic** mechanistic model
- Individual = farm

Infectious pressure

- At any time t, susceptible j has infectious pressure exerted on it by
    - all infected farms i
    - "Background" - eg wildlife

In a small time interval Δt:

$$P(j \text{ infected}) \approx T_j \cdot \Delta t$$

$$T_j = \beta_0 + \sum_{i \in \{I_i < I_j < N_i\}} \beta_{ij}(t) + \sum_{i \in \{N_i < I_j < R_i\}} \beta_{ij}^{\star}$$

IDReC
Infectious Disease Research Centre

The Model
Population structure

> ### Transmission Equations
>
> $$\beta_{ij} = q(i;\zeta)s(j;\xi)\left\{ K(i,j;\psi) + \mathbf{r}_{ij}^T \mathbf{p} + \mathbf{c}_{ij}^T \boldsymbol{\beta} \right\} \qquad i \in \mathbf{I}, j \in \mathbf{S}$$
>
> $$\beta_{ij}^* = \gamma q(i;\zeta)s(j;\xi)K(i,j;\psi) \qquad\qquad\qquad i \in \mathbf{N}, j \in \mathbf{S}$$

$q(i;\zeta)$ infectivity of i $s(j;\xi)$ susceptibility of j

$\mathbf{c}_{ij}$ static contact network $\mathbf{r}_{ij}$ dynamic contact network

$K(i,j;\psi)$ distance kernel ν "quarantine" parameter

The Model
Infection times

- Infection times are not directly observed

- Notification time is observed
  - Assume a distribution $f_D(\cdot)$ for Infection to Notification time
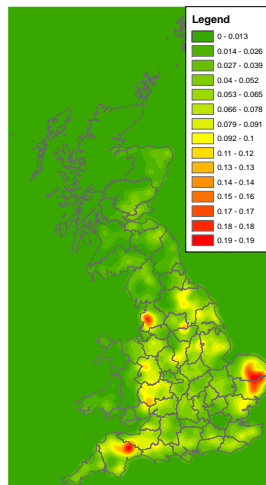 - eg Gamma(a, b)
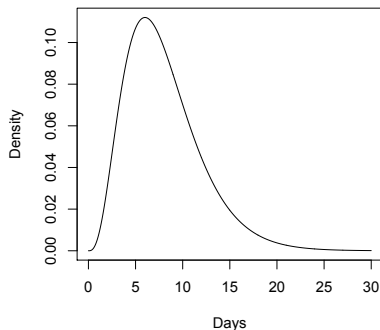
Approach to inference

- Construct a likelihood describing the continuous-time stochastic epidemic, *conditional on the infection times*

- Bayesian approach allows:
  1. Coherent inclusion of Prior information
     - Expert opinion
     - Previous disease outbreaks
  2. Natural framework to include unobserved data by data augmentation MCMC methodology
     - Unobserved infection times
     - Occult infections

IDReC
Infectious Disease Research Centre

Prior distributions

- Gamma for rates ($\beta \geq 0$)
- Beta for probabilities ($0 \leq p \leq 1$)
- Uniform for infection times ($-\infty \leq I < T_{obs}$)
- Uniform for occult status (0 or 1)

- Priors chosen to agree with expert opinion and previous knowledge of epidemics

Influenza H5N1 in British Poultry

- Extract from Great Britain Poultry Register (May 2006)
- 8363 registered poulty premises after data cleaning
    - Production stock only (10 types)

- 3 contact networks identified
    - Feed lorries
    - Slaughterhouse lorries
    - Company association

- OS National Grid coordinates for each premises

Legend
- 0 - 0.013
- 0.014 - 0.026
- 0.027 - 0.039
- 0.04 - 0.052
- 0.053 - 0.065
- 0.066 - 0.078
- 0.079 - 0.091
- 0.092 - 0.1
- 0.11 - 0.12
- 0.13 - 0.13
- 0.14 - 0.14
- 0.15 - 0.16
- 0.17 - 0.17
- 0.18 - 0.18
- 0.19 - 0.19

Infectious Disease Research Centre

The Model
Infection rate

Inter-farm transmission rate

$$\beta_{ij} = \eta_{sp,j} \left(\beta_1 C_{ij}^{FM} + \beta_2 C_{ij}^{SH} + \beta_3 C_{ij}^{CP} + \beta_4 e^{-\beta_6 \cdot \rho[i,j]} \right) \quad i \in \mathbf{I}, j \in \mathbf{S}$$

$$\beta_{ij}^{\star} = \eta_{sp,j} \left(\beta_5 e^{-\beta_6 \cdot \rho[i,j]} \right) \qquad\qquad\qquad\qquad\quad i \in \mathbf{N}, j \in \mathbf{S}$$
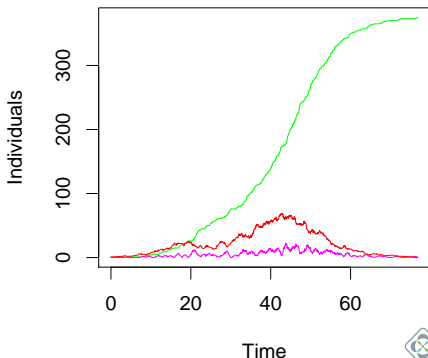
Time to Notification

$$P(D > d) = e^{\{-a(e^{\{b \cdot d\}} - 1)\}}$$

Time between infection and notification, $D|I = N - I$
where $a > 0, b > 0$.

Simulated epidemic

- No HPAI epidemic in the UK yet!

- Simulate epidemic on our dataset

| Time/days	Infections
0	1
14	10
25	61
50	290
76	375

Risk Prediction

Points of interest:

- How does the **parameter uncertainty** change with the amount of available data?

- How does the **risk** to specific farms change over the course of the epidemic?
  - Achieved by forward simulation of the epidemic using random sample of parameter values from posterior

- What are the locations of occult infections?

- Which farms would present the greatest danger to the population **if they were to be infected**?

IDReC
Infectious Disease Research Centre
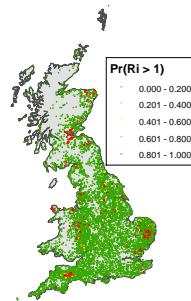
Farm-specific "R_0": R_i

"The expected number of farms each farm i might infect were it to be infected (in a totally susceptible population)."
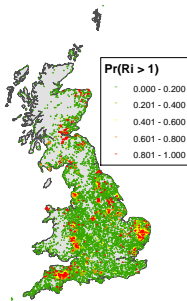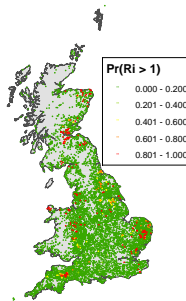
14 days   25 days   50 days

Table: Median R_i

bTB in Possums

Questions:

- How do we know where a possum lives?

- What determines possum contact networks?
    - Contact frequency?
    - Duration of contact?
    - Age, sex , etc?

- How do we infer a possum contact network?
    - $\longrightarrow$ feed into epidemic model!

The Ornstein-Uhlenbeck Process
SDE for mean-regressive diffusion process, eg Blackwell, 2008; Ramsey & Efford, 2010

- Possums move according to Brownian motion
- Should have a centre of attraction, eg. den, feeding sites, "average location"
- Assume independence and isotropy

Ornstein-Uhlenbeck Diffusion

$$dX_t = \alpha(\mu - X_t)dt + \beta dB_t$$

X_t – position at time t $\quad$ μ – "mean" location
α – drift coefficient $\quad$ β – diffusion coefficient
dB_t – Brownian motion

- Work with Alex James, Otago

IDReC
Infectious Disease Research Centre

The Ornstein-Uhlenbeck Process
Solution

- Solution of OU process is Gaussian!

> ### Solution of OU process
>
> $$X_{i,t+\delta}|X_{i,t} \quad \sim \quad \mathsf{N}\left(\mu_i + (x_{i,s} - \mu_i)e^{-\alpha_i\delta}, \frac{\beta_i^2}{2\alpha_i}\left(1 - e^{-2\alpha_i\delta}\right)\right)$$
>
> $$\overset{D}{=} \quad \lim_{\delta\to\infty} \mathsf{N}\left(\mu_i, \frac{\beta_i^2}{2\alpha i}\right)$$

The Ornstein-Uhlenbeck Process
Remarks

- Does not rely on coherent observations wrt time
- Home range $\rightarrow$ define as 2-d Gaussian distribution!
    - May allow validation by investigation of denning sites (eg. radio tracking)

- μ_i are nuisance parameters – Bayesian approach integrates this out

- "Overlap" between re-constructed limiting distributions
    - Proportion of time i is within distance r of j,

$$T_{ij}(r) = Pr(\|x_i - x_j\| < r)$$

   - CDF of difference between two bivariate Normal rv's.
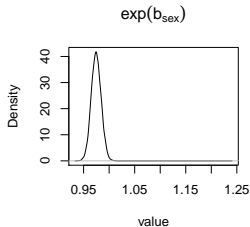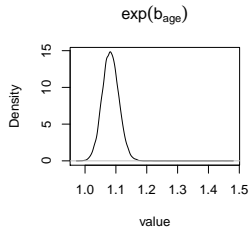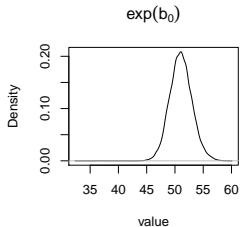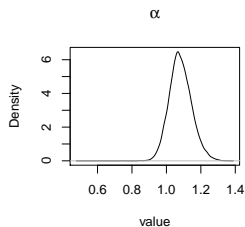
Possums in the Orongorongo valley

ORR valley model

$$\log \beta_i = b_0 + \beta_1 \cdot \text{sex}_i + \beta_2 \cdot \text{maturity}_i + \beta_3 \cdot \text{sex}_i \cdot \text{maturity}_i$$

- 4 grids, 360m x 360m: results Grid A only
- Capture-recapture data, non-coherent

- What is the possum home range?
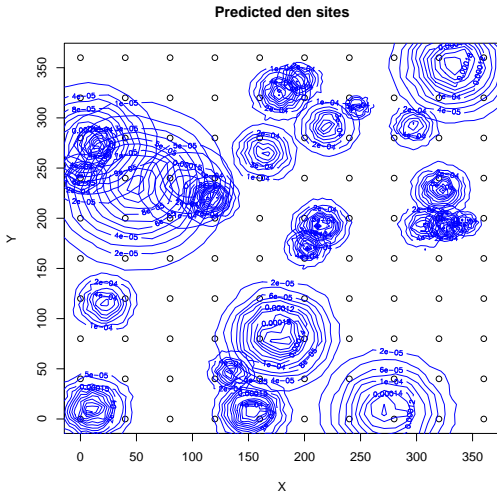- Does maturity (age) and sex influence home range?

- Priors non-informative

Possums in the Orongorongo valley
Coefficient marginal posteriors

- Slight $+$ve effect of age

- Slight evidence for females $\downarrow$ home range

Possums in the Orongorongo valley
"Centre of attration" posterior

Predicted den sites

Possums in the Orongorongo valley
Bayesian posterior prediction: possum home range

Posterior predictive home range

Next steps...

- OU process looks promising for modelling possum movements
- From our data, small effect of maturity/sex

- Grid A only – do we see the same effect in other grids?

- Couple possum OU processes using proximity logger data
  - We know possums were in proximity, but must impute location of contact

- Integrate into SIR-type epidemic model

Challenges

- Partial observation of the population

- Case detection
  - Coordinated with trapping times
  - Sensitivity and Specificity
  - Censored observations

- Computational aspects:
  - High dimensional missing data
  - DA-MCMC implementation
  - Non-centered parameterisation

- Validation of predictive distributions of epidemics

Acknowledgments

- University of Otago
  - Alex James

- Landcare
  - Dan Tompkins

- Massey University
  - Nigel French
  - Bryce Buddle
  - Kyle Richardson

- Marsden Fund