The Importance of being Uncertain:
quantification of uncertainty in epidemic prediction

Chris P Jewell and Gareth O Roberts
[Dept Maths and Stats, Lancaster University, UK]

Background

Mathematical modelling is common place in the study of epidemic diseases. Although it has been shown that such approaches can be used to make qualitative statements about the effect of control strategies on the likely outcome of an epidemic, their main weakness is in a lack of rigorously estimated parameters. This poster outlines a Bayesian methodology for making inference on parameters from data in order to make simulation models more quantitatively predictive.

What is uncertainty?

***Uncertainty* is a reflection of confidence in a parameter taking a particular value.**

The amount of uncertainty depends on the amount of data available – sample size.

Frequentist: A parameter takes a point estimate, the mean, with a standard error (Normal assumption)

Bayesian: Represents a parameter as a probability distribution.

The Approach

AIM: To use field data from an epidemic to capture in ***real time*** the values of parameters in a ***specified epidemic model***. These are then used to simulate from *the same model* to provide a *quantitative* risk-prediction at any point during an epidemic.

• Fully ***Bayesian*** approach

• The epidemic is modelled at the ***individual*** (eg farm) level

• **SINR** model:

➢ Individuals can be *susceptible*, *infected*, *detected*, and *removed*

S → I → N → R

• **Covariates** are obtained on a farm-level basis during peacetime. This allows us to study the relationships between farms in space, and commerce.

• **Epidemic** data is taken from the field *during* a disease outbreak. For each farm, we have a pair of **detection** and **notification** times, together with any contact tracing data that might be available.

Expert Opinion

• Obtained by talking to industry experts, analysing data from previous epidemics in the same or other species.
• Incorporated as Bayesian *prior information*

Data

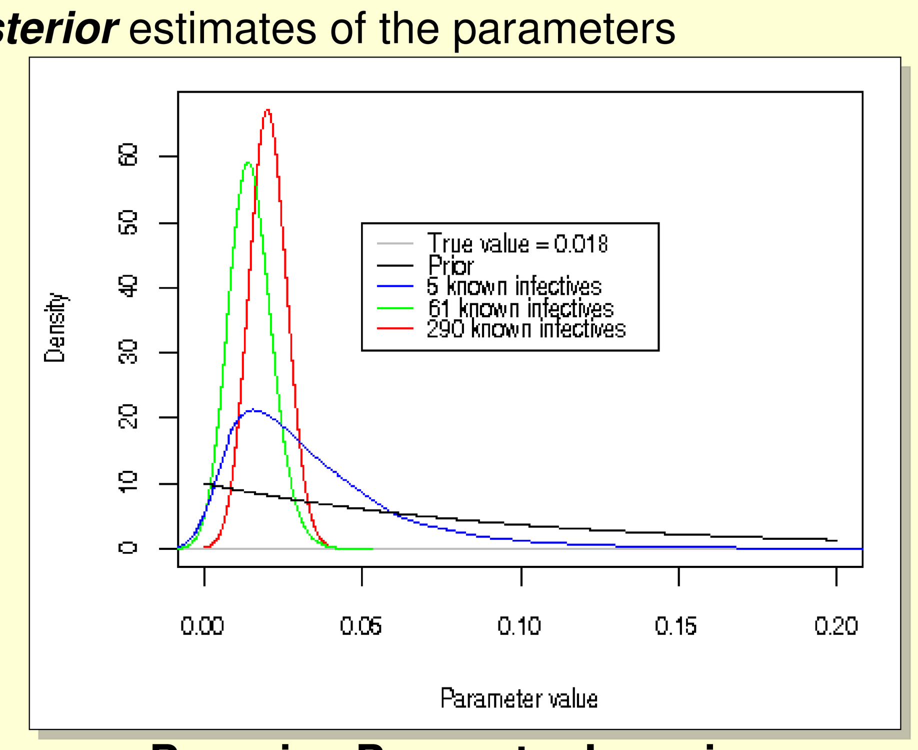
Parameter Inference

Reversible Jump Markov Chain Monte Carlo

• *Continuous time* inhomogeneous Poisson process for infection and notifications
• **Posterior distributions** of the parameters of interest are obtained
• **Missing data**:
  ➢ Unobserved infection times
  ➢ Presence of undetected infections in the population.

Parameter Estimates (with uncertainty)

• **Posterior distributions** of the parameters of interest are obtained
• At the start of the epidemic, there is little field data and the posterior estimates are highly influenced by the ***prior*** information
• As the epidemic progresses, we ***learn*** from field data which ***updates*** the prior information to give ***posterior*** estimates of the parameters



Bayesian Parameter Learning

Simulation

• Stochastic simulations are then used to *predict* the future course of the epidemic:
  ➢ ***Same model*** as parameter inference
  ➢ Many simulations are used, with parameters ***drawn from the posterior distributions***
  ➢ This naturally incorporates the ***uncertainty*** in the parameter estimates

Prediction (with uncertainty)

Why uncertainty?

***Quantitative* risk predictions for epidemics require simulations driven by *formally* estimated parameters.**

Threshold theorem: (homogeneous large population)

If $R_0 < 1$ – epidemic dies out with probability 1

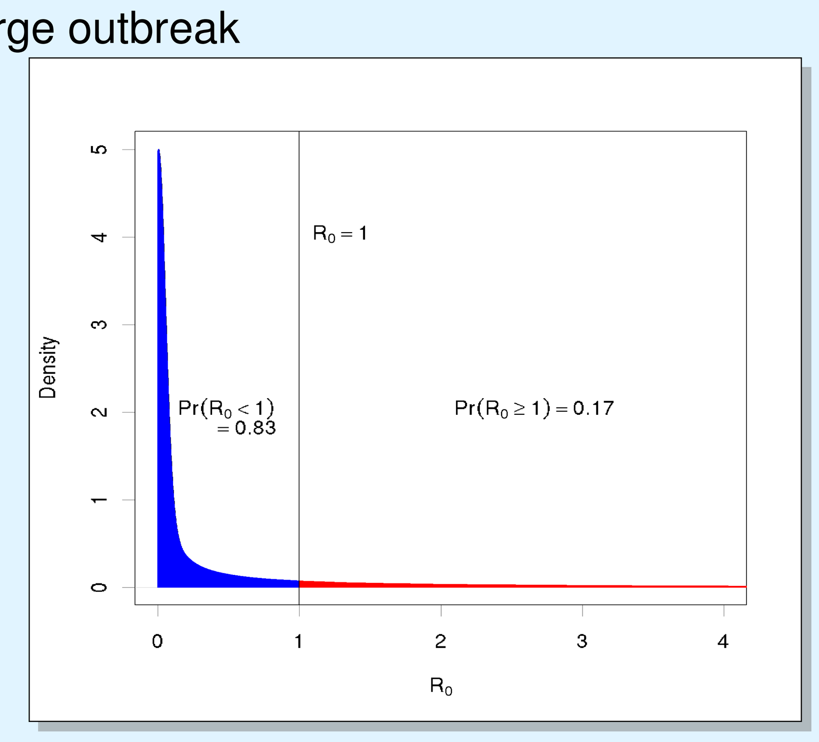
If $R_0 \geq 1$ – epidemic able to take hold and dies out with probability less than 1

Example:

R_0 is estimated to be 0.999

• If a point value is used, we predict that the epidemic will not take off.

• Introducing uncertainty allows R_0 to exceed 1 with some probability, increasing the chance of a large outbreak



Introducing uncertainty into R_0 – the density plot shows that R_0 has a significant probability of being greater than 1

Quantities such as:

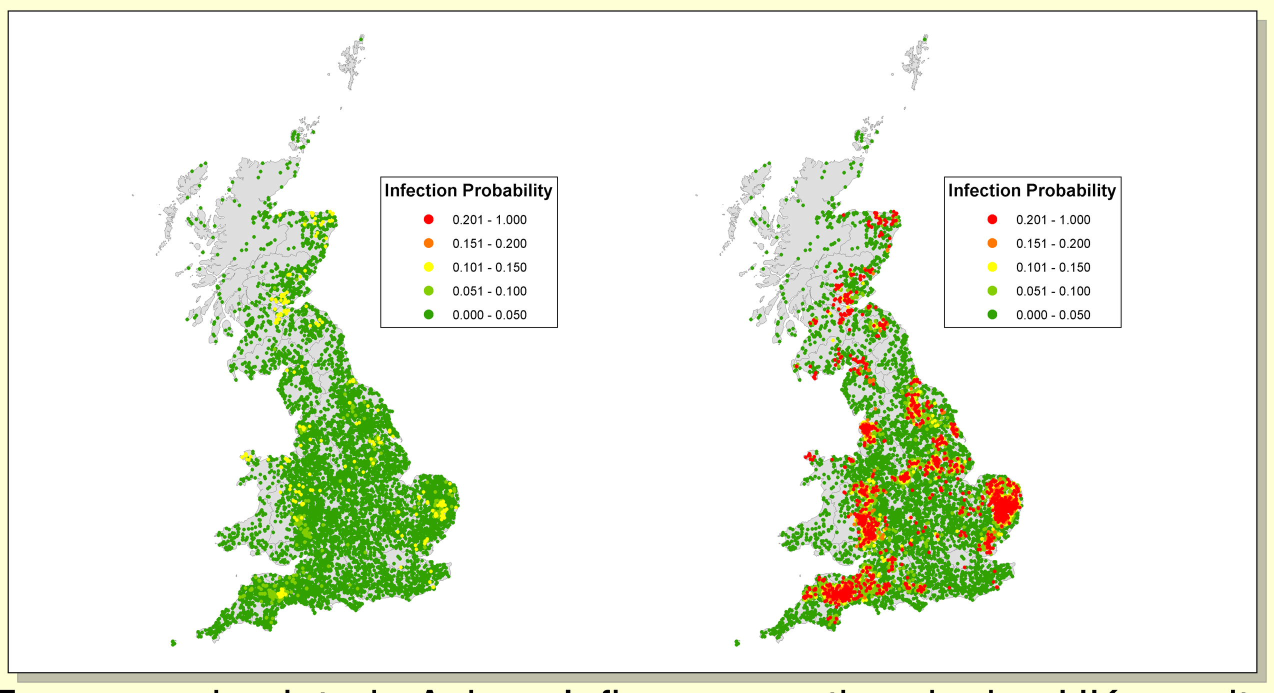
• Probabilities of individuals becoming infected (right)

• Individual specific R – the danger that an individual would pose to the population if it were infected

• The presence and spatial location of *undetected* infections

• Who started the epidemic?

• **Quantification of uncertainty** on the above



From a simulated Avian Influenza outbreak in UK poultry: ***probability of farms becoming infected*** at day 10 given the current epidemic: without uncertainty (left), with uncertainty (right)

References

Gibson GJ, Renshaw E (1998) Estimating parameters in stochastic compartmental models using Markov Chain methods. *IMA J. Math. Appl. Med. Biol.*, 15:19

O'Neill PD, Roberts GO (1999) Bayesian Inference for Partially Observed Stochastic Epidemics. *J. Roy. Statist. Soc. Ser. A*, 162:121

Britton T, O'Neill PD (2002) Bayesian inference for stochastic epidemics in populations with random social structure. *Scand. J. Statist.*, 29:375

Neal PJ, Roberts GO (2004) Statistical inference and model selection for the 1861 Hagelloch measles epidemic. *Biostatistics*, 5:249

This work was supported by research grant VTRI VT0103 from the higher Education Funding Council for England and the Department for Environment, Food and Rural Affairs.

Thanks to Theo Kypraios (University of Nottingham) and Rob Christley (University of Liverpool) for their invaluable help in this work.

Contact: c.jewell@lancaster.ac.uk