

Statistical Inference in a Zombie Outbreak Model: Is it Safe to Go Out Yet?

Des Higham

Department of Mathematics and Statistics
University of Strathclyde, Glasgow, Scotland

Uncertainty Quantification Workshop
ICMS, Edinburgh
May, 2010



Motivation

- Illustrate the Bayesian framework for parameter estimation in ODEs
- Contrast with least squares “common sense” approach
- Show state-of-the-art computations
- Use an eye-catching, if somewhat frivolous, application

P. Munz, I. Hudea, J. Imad, R. Smith?

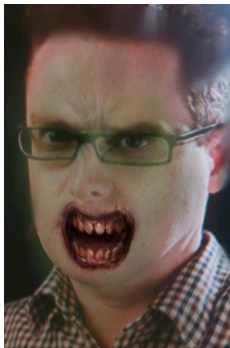
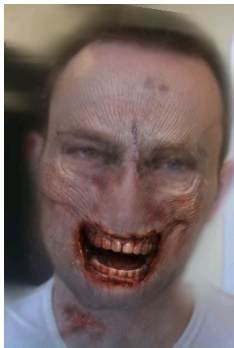
in **Infectious Disease Modelling Research Progress**,
Eds J. Tchenche and C. Chiyaka, **2009**

B. Calderhead, M. Girolami, D. J. Higham

Technical Report 6, 2010, U. Strathclyde

B. Calderhead, M. Girolami, D. J. Higham

Technical Report 6, 2010, U. Strathclyde



ODE Model

In terms of

- susceptibles (humans), $S(t)$
- zombies, $Z(t)$
- removed ('dead' zombies), $R(t)$ —these may return as zombies

$$S' = -\beta SZ$$

$$Z' = \beta SZ + \zeta R - \alpha SZ$$

$$R' = \alpha SZ - \zeta R$$

Given some observations, we wish to, e.g.,

- infer values for the model parameters
- predict future behaviour of the system

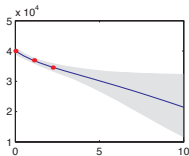
Bayes' Theorem

$$P(\theta|\mathbf{Y}, M) = \frac{P(\mathbf{Y}|\theta, M)P(\theta|M)}{P(\mathbf{Y}|M)}$$

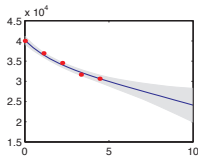
$$P(\theta|\mathbf{Y}, M) \propto P(\mathbf{Y}|\theta, M)P(\theta|M)$$

Posterior Outputs

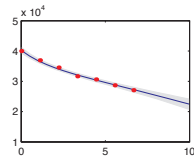
3 Data Points



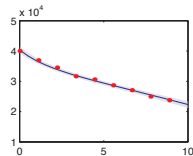
5 Data Points



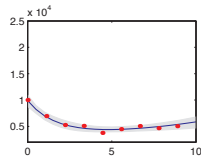
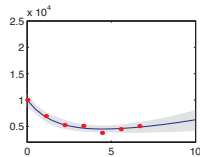
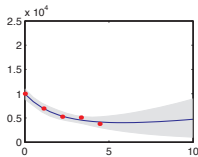
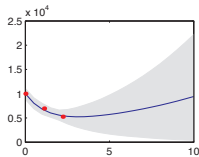
7 Data Points



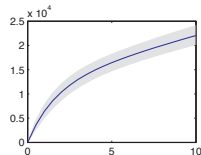
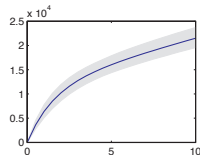
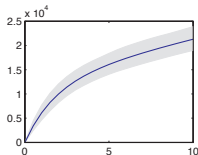
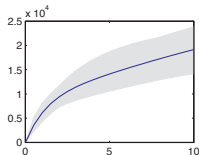
9 Data Points



Human Population (Observed)

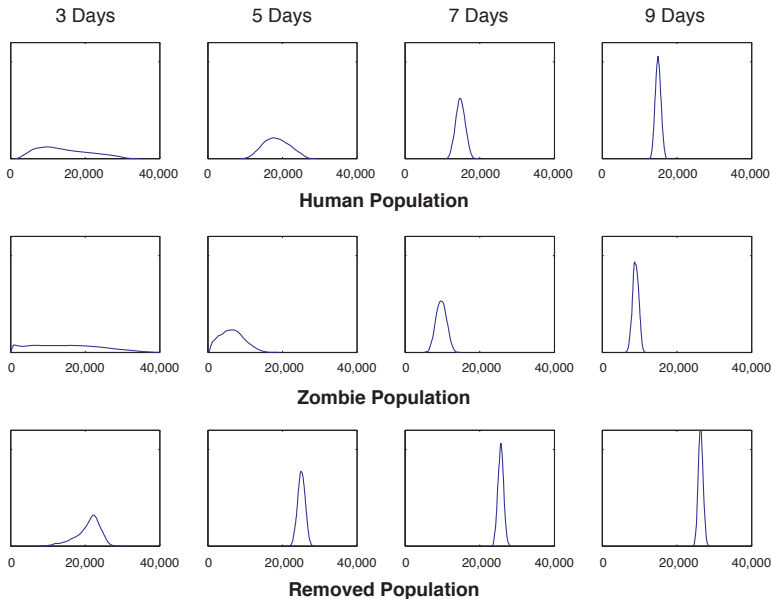


Zombie Population (Observed)



Removed Population (Unobserved)

Predicted Levels at Day 15



Competing Model?

Zombies only attack in pairs

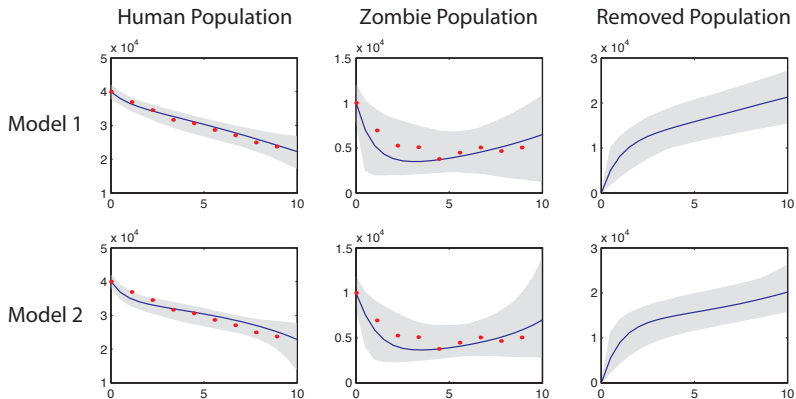
$$S' = -\beta SZ^2$$

$$Z' = \beta SZ^2 + \zeta R - \alpha SZ$$

$$R' = \alpha SZ - \zeta R$$

Model comparison is possible with Bayes factors

Competing Posterior Outputs



Bayes factor for Model 1 versus Model 2 is ≈ 600

Decisive evidence in favour of Model 1

Pros and Cons

- Need to specify **priors** (**log barrier function**)
- Need to choose an error model in order to construct **likelihood** (**objective function**)
- Need to **search** and then **integrate** over high dimensional space (**optimize**)

But, we can then quantify uncertainty :

- **posterior distribution for model parameters**
- **posterior model output**
- **systematic model comparison**