Is It Safe To Go Out Yet? Statistical Inference in a Zombie Outbreak Model*

Ben Calderhead† Mark Girolami‡ Desmond J. Higham§

June 14, 2010

1 Mathematical Modelling with Ordinary Differential Equations

Ordinary differential equations (ODEs)—the kind that we meet in introductory calculus classes—have proved to be extremely useful tools for describing, in quantitative terms, how physical systems change over time. Some systems are so well understood that they follow widely-accepted "laws of motion", for example planetary orbits, ballistic missiles, elastic springs, chemical reactions and radioactive decay. In other cases, where there may be a range of complicated objects interacting in ways that are impossible to pin down, ODEs can still be extremely effective for characterizing the main features of the system. *Modelling* with ODEs, that is, constructing ODEs that have some explanatory and predictive power, involves identifying the essential features of a physical system and converting these into a descriptive mathematical framework. Success stories abound throughout the physical and engineering sciences, and, of course, the world of science fiction [9].

In the case of zombie outbreaks, just as in any other modelling context, a good model allows us to

explain the key mechanisms at work and how they interact,

predict the future state of the system and, more excitingly,

*Prepared for the proposed book *Academics on Zombies*, edited by R. Smith?

†Department of Computing Science, University of Glasgow.
‡Department of Computing Science, University of Glasgow.
§Department of Mathematics and Statistics, University of Strathclyde.

theoretically and computationally investigate "what-if" scenarios that may be out of the reach of experimental scientists because of technological or budgetary limitations, or ethical considerations.

To be specific, an accurate zombie outbreak model could give *quantitative* feedback to frightened individuals with concerns such as

- Soldiers patrolling the area have reported daily observed zombie numbers for the past 5 days as $123, 127, 104, 92, 74$. Is it safe to go out yet?
- If I meet a zombie, what is my chance of fighting it off?

It could also inform brave but under-resourced action heroes who need to make tough decisions such as

- How many soldiers should we mobilize, and how many of these will survive?
- What scale of quarantine would be worthwhile?
- How effective does this cure need to be?

However, just as in any realistic modelling context, the zombie outbreak ODE models of [9] involve parameters whose values are not immediately obvious. If we picture the model as a "black box," then there will be dials on the box that must be tuned to the particular circumstances of the outbreak, such as the malevolence of the zombies and the pluckiness of the humans. While there may be a few general issues that can be addressed without detailed knowledge of these parameters, to make the mathematical model really useful, we must *calibrate* any unknown parameters using observed data.

Recently, the challenge of parameter estimation for ODE models has attracted the attention of the statistical inference community, and some powerful tools have been developed that go far beyond the traditional "least-squares" style point estimates [3, 11, 12]. The zombie scenario therefore gives us the opportunity to raise the profile of this emerging field, and point out the advantages that arise when ideas from applied/computational mathematics and Bayesian reasoning are brought together.

Using models motivated by [9], we will consider how much prior knowledge and how much observational data is needed in order to make useful inferences about a zombie outbreak. We will also introduce the *model selection* issue [11]. Suppose we hear incompatible rumors that

(a) zombies only attack alone,

(b) zombies only attack in pairs.

The two scenarios correspond to different ODE models, so to investigate which rumor is more likely to be true, we could ask which of the two models best explains the observed data. This requires us *simultaneously to calibrate and compare two or more ODE models*; a topic that has only recently been tackled systematically [2, 11].

2 Simple Model

In this section we illustrate some key ideas by focusing on a very simple zombie outbreak model. Imagine a population consisting only of humans and zombies. We will let $S(t)$ and $Z(t)$ represent the number of humans and zombies at time t. In fact, $S(t)$ and $Z(t)$ will take real values, so it is more appropriate to think of these values as the concentration levels of the two species. We suppose that there is just one event that can cause a change in the levels: a zombie may attack a human and, if successful, convert that human into a zombie. We note immediately that in this simple world the population level of humans is forced to decrease to zero and the population level of zombies will correspondingly increase until all the humans are used up. This modelling assumption requires us to introduce a single parameter, a rate constant, β, that characterizes the ability of the zombies to find and infect humans. The larger the value of β, the more virulent the zombies. Simple *mass action* modelling then leads us to the ODE

$$S'(t) = -\beta S(t)Z(t), \tag{1}$$
$$Z'(t) = \beta S(t)Z(t). \tag{2}$$

Here S' and Z' denote the time derivatives, that is

$$S'(t) \equiv \frac{dS(t)}{dt} \quad \text{and} \quad Z'(t) \equiv \frac{dZ(t)}{dt}.$$

So the ODE system (1)–(2) specifies the rate of change of the two population levels. Adding the two equations, we see that $(S(t) + Z(t))' = 0$, so the rate of change of the total population, humans plus zombies, is zero. This is intuitively obvious, each time we lose a human we gain a corresponding zombie. So $S(t) + Z(t)$ remains constant for all time. Letting the constant K denote this overall population size, we have $S(t) + Z(t) = K$. We may then replace $Z(t)$ by $K - S(t)$ in (1) to get a single ODE

$$S'(t) = -\beta S(t)\left(K - S(t)\right). \tag{3}$$

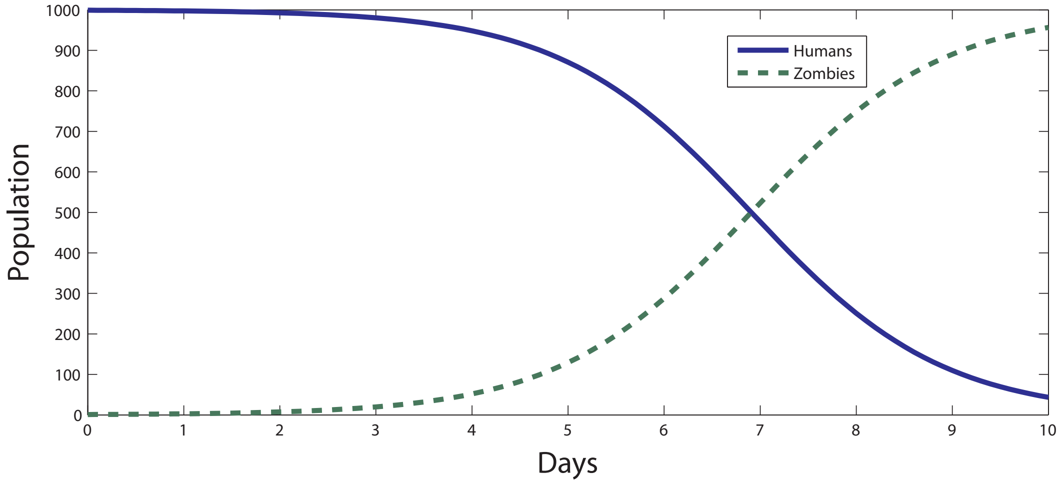

Figure 1: Simple zombie model with parameter $\beta = 0.001$.

This ODE fits into the class of *logistic equations*. Although nonlinear, it is sufficiently simple to admit a pencil-and-paper solution. This may be written

$$S(t) = \frac{S(0)K}{S(0) + (K - S(0))\, e^{\beta K t}}. \tag{4}$$

Here $t = 0$ represents the initial time when we start to monitor the populations, so $S(0)$ is the initial level of the human population.

Let us first assume that the size of the joint population, K, of humans and zombies is known, and that the initial number of humans, $S(0)$, is also known. This leaves us with one unknown parameter, β. This parameter may be interpreted as the rate at which humans are converted into zombies and is measured "per zombie per day". Therefore if initially there are 1000 humans, a rate of $\beta = 0.001$ initially corresponds to $S(0) \times \beta = 1000 \times 0.001 = 1$ human being converted by each zombie every day. Note that the number of humans being converted depends on both the number of humans and the number of zombies at any particular time, and by solving our differential equations we may see how these two populations evolve relative to one another.

Figure 1 shows how a population of 1000 humans diminishes to less than 10 after just 10 days, based on a single zombie initially attacking and converting humans at a rate of one human per zombie per day. We may see the effect of doubling the rate constant to $\beta = 0.002$ in Figure 2. In this case the population dwindles to less than 10 after just 5 days, and by the 7th day the humans have effectively died out.

These predictions have been made under the assumption that we know the exact rate at which zombies attack and convert humans. Given an initial

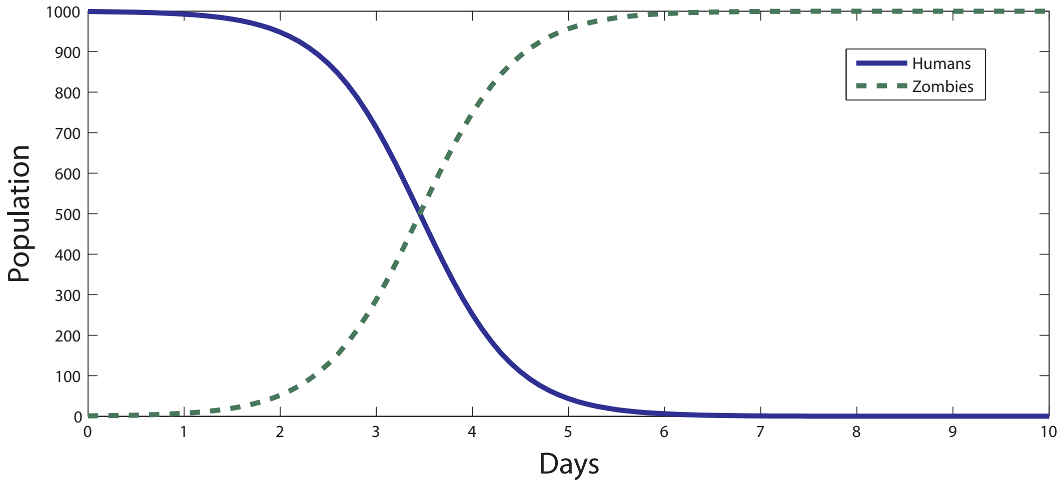

Figure 2: Simple zombie model with parameter $\beta = 0.002$.

number of humans and zombies, along with a rate constant, our model tells us how the populations evolve. In a more realistic scenario, we would not know the rate constant, but instead have (usually inaccurate) observations of how the human population changes over a period of time. The goal in this case is then reversed: given some observations regarding the number of humans at certain time points, we want to estimate the rate constant β such that our model best describes the situation as we see it. Indeed, estimating these rate constants is often very challenging and is known as the *inverse problem*. Once we have inferred the rate parameter β, we may then quantify the likelihood of future scenarios, based on the knowledge that our model adequately describes the past.

It may be argued that there is no single "true" value of β because the model, which is based on simplifying assumptions, does not capture every detail of the physical system. More fundamentally, there will also be measurement errors and uncertainties in the data—in the case of a zombie attack, humans in hiding may go unrecorded and zombies may lurk unnoticed in dark corners. Therefore, there may well be many values of β that are approximately as good as each other at describing the data. It makes sense then to determine the probability distribution over the *most likely* values for β as opposed to a single "best" value. Statistical inference, and in particular *Bayes' theorem*[1], gives us a mathematical framework in which to carry out these calculations in terms of probability distributions. We refer to [10]

[1]The Reverend Thomas Bayes was a British mathematician and Presbyterian minister, born at the start of the 18th century. He developed a theory for calculating inverse probabilities. The more general formulation in use today subsequently bears his name.

for an accessible introduction to the Bayesian framework and [7] for a more comprehensive treatment.

There are three levels of inference that we may ultimately wish to carry out. The first determines the parameters with which the model plausibly describes the data. This is the probability of the free parameters $\boldsymbol{\theta} = [\theta_1, ..., \theta_n]$ given some data $\mathbf{Y}$ and a particular model M, which may be written as $P(\boldsymbol{\theta}|\mathbf{Y}, M)$. In our example there is just one parameter, $\boldsymbol{\theta} = \beta$, and $\mathbf{Y}$ is a vector of observations at a number of time points. The second level of inference sheds light on the uncertainty associated with our choice of model, and this is the probability of a particular model M given the data $\mathbf{Y}$, written as $P(M|\mathbf{Y})$. Finally, the third level of inference describes the probability of a *prediction* given the data, and this prediction may be based on multiple plausible models which are weighted according to their relative probabilities. In order to estimate these probability distributions we can make use of Bayes' theorem, which gives us a method of combining prior knowledge and newly obtained data in a mathematically consistent manner. Key to the Bayesian approach is the idea that observations are inherently uncertain, so that a single datapoint is assumed to be but one sample from some underlying probability distribution, which is represented as the likelihood of the data given a model and its current set of parameters, $P(\mathbf{Y}|\boldsymbol{\theta}, M)$. The prior distribution, $P(\boldsymbol{\theta}|M)$, which is often simply referred to as "the prior," characterises our initial knowledge or belief regarding plausible values of the parameters.

Let us therefore first define our prior and likelihood distributions for the model describing simple human/zombie interactions. A prior can be constructed by considering a reasonable time scale for the process; we may argue that all the action will not be over in a single day. This gives an upper limit of 1 for the rate constant β—with that value $S(0) \times \beta = S(0)$, so all $S(0)$ humans could be converted by one zombie on the very first day. Likewise, a reasonable minimum rate is $\beta = 0$, in which case zombies never convert humans. With no further information, it is reasonable to take the view that, a priori, any rate constant between 0 and 1 is equally likely. So we may choose our prior on β to have a uniform distribution over this range.

The likelihood is a measure of the goodness of fit between the data and the output of the model. The choice of which probability distribution to employ should depend on the problem context. For some modelling scenarios, where for example the observed data is the number of counts occurring within a particular time interval, the choice of a Poisson distribution is appropriate. Alternatively, if the observed data is obtained from estimates which may be affected by a large number of small unknown random factors, then due to the the Central Limit Theorem [7, 10] the associated error may be well approximated by a Gaussian distribution. In the case of modelling our zombie attack, we shall assume that the estimated population levels are subject to small unknown errors, with the final estimates combining local intelligence, large numbers of individual sightings and so on. We also assume that the errors at different observation times are independent. We therefore define the likelihood function to be the following quantitative measure of the agreement between the model output and the observed data over all time points

$$L = P(\mathbf{Y}|\boldsymbol{\theta}, M) = \prod_t N_{Y(t)}(S(t), \sigma^2). \tag{5}$$

Here,

- $P(\mathbf{Y}|\boldsymbol{\theta}, M)$ represents the probability of an observation given a model M with parameters $\boldsymbol{\theta}$,
- $Y(t)$ is the observation at time t,
- $S(t)$ is the output of the model at time t, given parameters $\boldsymbol{\theta}$,
- $N_x(\mu, \sigma^2)$ is the density function for a Gaussian with mean μ and variance σ^2, that is,

$$N_x(\mu, \sigma^2) := \frac{\exp(-(x-\mu)^2/(2\sigma^2))}{\sqrt{2\pi\sigma^2}}.$$

The variance σ^2, which may be regarded as the inherent level of uncertainty in the data, could either be estimated and fixed in advance, or inferred along with the other parameters. We note that the product (5) corresponds to an overall normal distribution. In the case of our simple zombie model, $S(t)$ is easily calculated from (4). For more complex models however it will be necessary to compute a numerical solution for the ODE model.

In summary, for a particular combination of model parameters and initial condition, given an observation of the number of surviving humans at a known point in time, we compute the likelihood by taking a Gaussian density centered on the model prediction and then finding the value of the density function at the observed value. We repeat this for each data point and multiply the answers together.

Bayes' theorem allows us to update our initial belief about the parameter values, as defined by our prior, by taking the data into account. Our updated knowledge is then quantified by the posterior distribution $P(\boldsymbol{\theta}|\mathbf{Y}, M)$ (first level of inference) by combining our prior distribution with the likelihood function in the following mathematically consistent manner,

$$\begin{aligned} P(\boldsymbol{\theta}|\mathbf{Y}, M) &= \frac{P(\mathbf{Y}|\boldsymbol{\theta}, M)P(\boldsymbol{\theta}|M)}{P(\mathbf{Y}|M)} \\ &\propto P(\mathbf{Y}|\boldsymbol{\theta}, M)P(\boldsymbol{\theta}|M). \end{aligned}$$

Noting that the marginal likelihood $P(\mathbf{Y}|M)$ is constant for a particular model M, we see that it may be calculated as the integral of the likelihood times the prior over all parameter values,

$$\begin{aligned} P(\mathbf{Y}|M) &= \int \ldots \int P(\mathbf{Y}, \boldsymbol{\theta}|M) d\theta_1 \ldots \theta_n \\ &= \int P(\mathbf{Y}|\boldsymbol{\theta}, M)P(\boldsymbol{\theta}|M) d\boldsymbol{\theta}. \end{aligned}$$

Performing the second and third levels of statistical inference over ODE models is challenging precisely because we have to estimate this integral, which is generally analytically intractable and high dimensional. Only recently has it been demonstrated that this integral may be efficiently and accurately estimated using a technique called thermodynamic integration [2], and we shall employ this later in the paper to discriminate between competing model hypotheses describing zombie populations and to quantify predictions of future behaviour.

Finally, we need a method of sampling from the posterior distribution, which in our case is analytically intractable, since we can't calculate the marginal likelihood analytically for our models based on nonlinear differential equations. Fortunately we can instead employ the Metropolis-Hastings algorithm to draw a series of random samples from an approximation of the posterior. This Markov chain Monte Carlo method was first developed by nuclear physicists at Los Alamos Laboratory, New Mexico in the 1950s, who were investigating methods of simulating the random behaviour of neutrons in the fissile material of atomic bombs [6]. Later it was realised that the method could be used more generally to simulate from virtually any probability distribution and a generalisation was provided by Hastings [5]. Modern computing power is now making this technique feasible across a tremendous variety of applications. In particular, in Bayesian statistics it allows us to draw accurate samples from the posterior distribution even if the marginal likelihood $P(\mathbf{Y}|M)$ is not known. The basic approach is relatively straightforward to implement; given a current value of our parameter β_c we can randomly generate a new state β_n from a *proposal distribution* $Q(\beta_n|\beta_c)$, which depends only on the current state. For efficiency, the proposal distribution

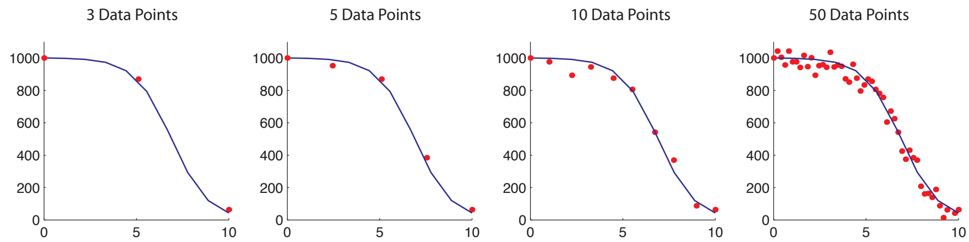

Figure 3: $[3, 5, 10, 50]$ data points generated from the simple zombie model over 10 days with parameter $\beta = 0.001$ and Gaussian distributed noise with a standard deviation of 50.

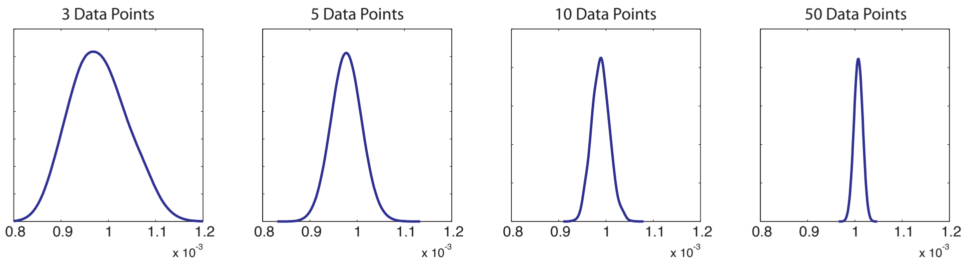

Figure 4: Posterior output from the simple zombie model with parameter $\beta = 0.001$ and Gaussian distributed noise with a standard deviation of 50, as shown in Figure 3. As the number of data points increases, so the posterior becomes more sharply peaked around the true value of β.

should be as similar as possible to the target distribution you wish to sample from — this ensures that we do not waste too much time proposing unlikely values. In practice, since we generally have little information about the posterior distribution in advance, it suffices to employ a Gaussian distribution with mean β_c and some variance which may be chosen so as to achieve an acceptance rate of between 20% and 40%. This new state is accepted with probability

$$\min\left\{\frac{P(\beta_n)Q(\beta_c|\beta_n)}{P(\beta_c)Q(\beta_n|\beta_c)}, 1\right\},$$

where $P(\beta)$ is the probability distribution we wish to sample from, which in our case is the posterior $P(\beta|\mathbf{Y}, M)$. Loosely, this technique allows us to search through the set of possible parameter values, spending most of our time in the "hot spots" where parameter values are most promising.

We are now ready to infer a posterior distribution over the parameter values given some data. We will do this using "artificial" data so that we can judge the performance of the algorithm under controlled circumstances.

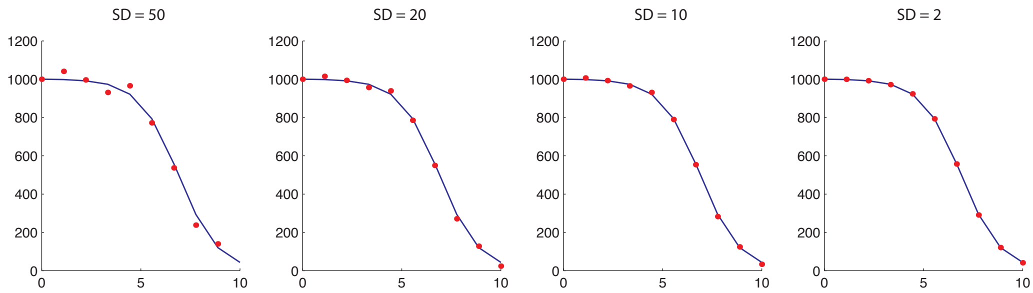

Figure 5: 10 data points generated from the simple zombie model over 10 days with parameter $\beta = 0.001$ and Gaussian distributed noise with a standard deviation of $[50, 20, 10, 2]$.

More precisely, we evaluate the solution (4) of the the differential equation for a chosen value of β at a number of time points and add some Gaussian distributed noise with known variance to the solution to generate some experimental data. We generated four data sets this way, as shown in Figure 3. We then treat this data as though it came from the model with an unknown value of β, and see what quality of inference is possible. Figure 4 shows how the posterior distribution over β becomes more sharply peaked around the true value as the number of data points increases from 3 to 50. The noise induces a noticeable bias when using just 3 data points, although the posterior is reasonably large at the true value of 0.001. Figures 5 and 6 show how, as one might expect, as we add less noise to the data, so the posterior distribution becomes less diffuse, indicating a greater confidence in the range of values for which the model could plausibly describe the data. We may also examine the effect of our prior on the posterior. Changing the prior from uniform over $[0, 1]$ to uniform over $[0, 0.01]$ has very little difference on the posterior of β, as shown in the left and middle pictures of Figure 7. If however we were to badly mis-specify the prior by, for example, setting it to be a sharply peaked Gaussian distribution over the wrong value, we observe a biased and rather skewed posterior distribution in the picture on the right in Figure 7. This type of mis-specified prior may often be diagnosed by comparing the prior and posterior.

Having shown how Bayesian inference can be applied in a very simple ODE setting, in the next section we will move on to a more realistic model where (a) there is more than one unknown parameter and (b) the ODE solution cannot be written down explicitly.

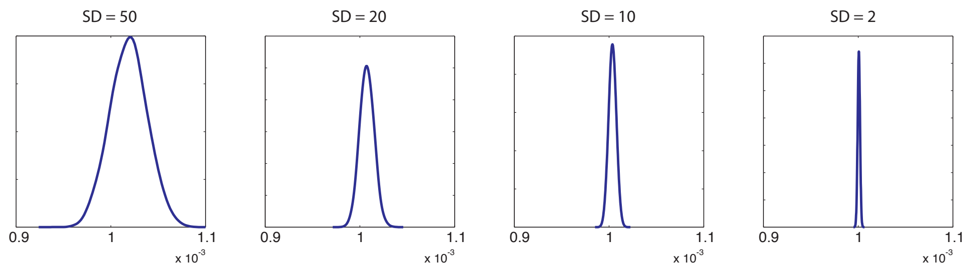

Figure 6: Posterior output from the simple zombie model with parameter $\beta = 0.001$ and 10 data points. As the standard deviation of the added noise decreases, so the posterior becomes more sharply peaked around the true value of β.

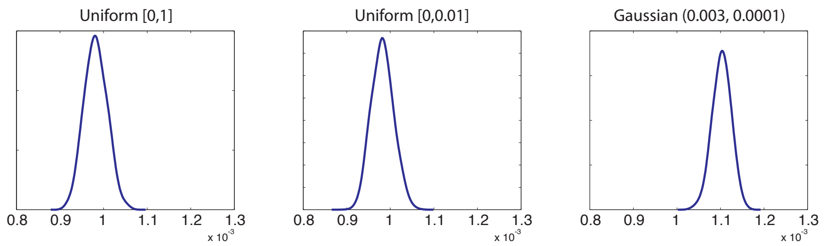

Figure 7: Posterior output from the simple zombie model with parameter $\beta = 0.001$, 3 data points and Gaussian noise with standard deviation 20. Changing the uniform prior from the range $[0, 1]$ to $[0, 0.01]$ has little effect on the posterior. However, a mis-specified Gaussian prior sharply peaked around 0.003 with standard deviation 0.0001 results in a biased posterior distribution based on this small amount of data.

3 More Realistic Model

The model derived in section 2 of [9] splits the overall population into three classes. At each time t we have

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

As in our simple model (1)–(2), humans, now called *susceptibles* to be consistent with the traditional population dynamics literature, are liable to attack by zombies, and such an attack may convert the human into a zombie. However, we now allow for the possibility that a human can prevail in a human-zombie encounter. In this case the human emerges unscathed and the zombie joins the removed class. The term *removed* is used rather than

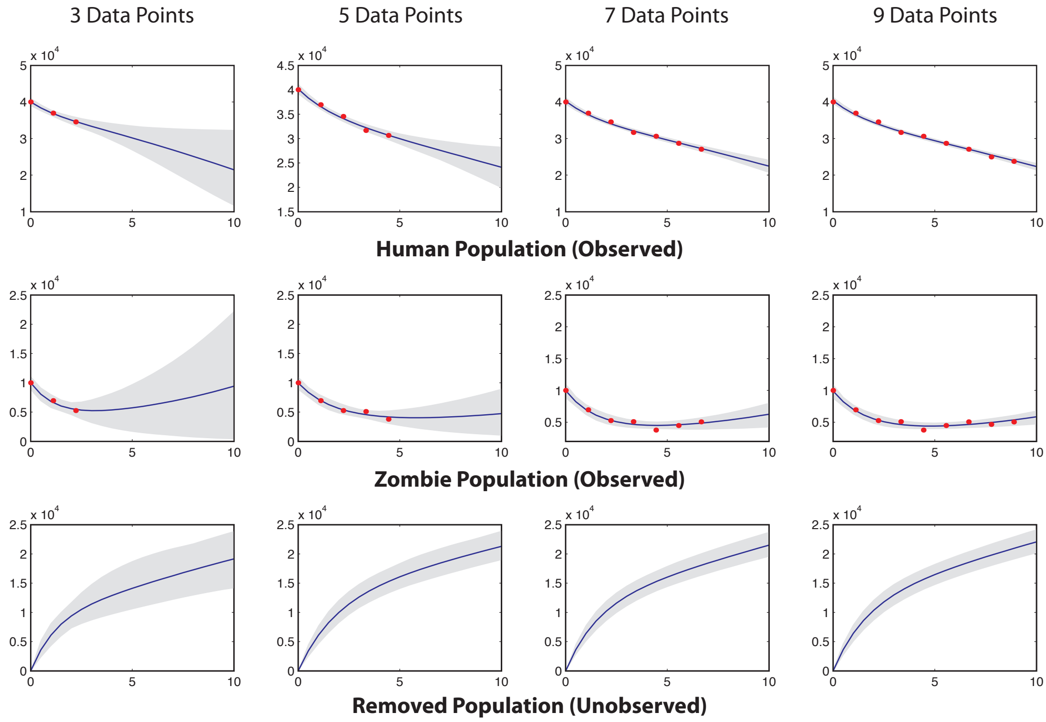

Figure 8: Posterior output from the complex zombie model (6)–(8) with parameters $\beta = 0.00001$, $\alpha = 0.00002$, $\zeta = 0.1$, and Gaussian distributed noise with a standard deviation of 500. As the number of data points increases, so the uncertainty in the posterior model output decreases.

dead because a human can never completely kill a zombie, at some later time the "dead zombie" may resurrect back to regular zombie status. These modelling assumptions lead us to the following ODE system from [9][2]

$$
\begin{aligned}
S' &= -\beta S Z, & (6)\\
Z' &= \beta S Z + \zeta R - \alpha S Z, & (7)\\
R' &= \alpha S Z - \zeta R. & (8)
\end{aligned}
$$

The (constant, positive) parameters in the model are:

- β: a rate constant for a susceptible/zombie encounter causing a susceptible to turn into a zombie. This parameter arose in the simple model of section 2. A larger value of β applies when zombification is more likely.

[2]For simplicity we consider only the short timescale regime ($\Pi = \delta = 0$ in [9]), so no new humans are born and no humans die from any other causes.

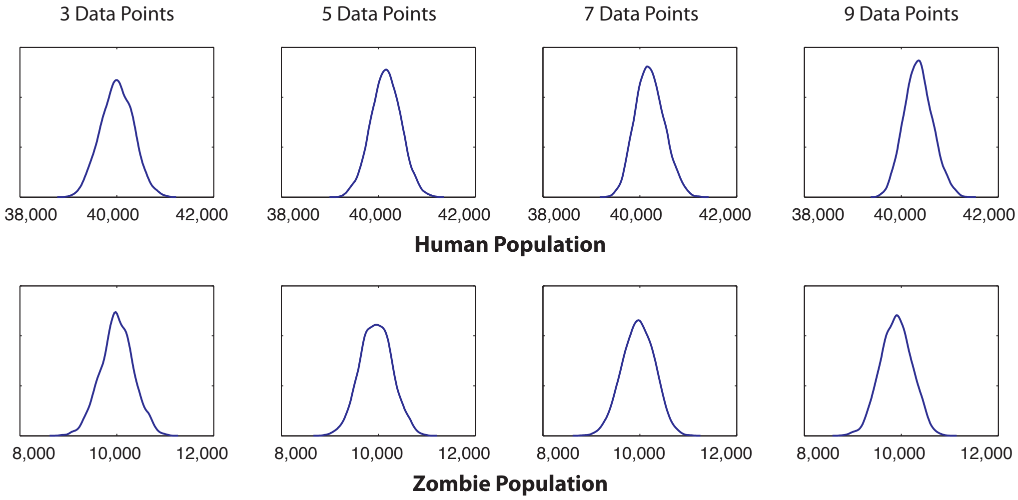

Figure 9: Posterior distributions of the inferred initial conditions from data generated by the complex zombie model with parameters $\beta = 0.00001$, $\alpha = 0.00002$, $\zeta = 0.1$, and Gaussian distributed noise with a standard deviation of 500. We observe that the initial conditions are in this case relatively insensitive to the number of data points observed.

- α: a rate constant for a susceptible/zombie encounter causing a zombie to become removed. A larger value of α applies when removal is more likely.

- ζ: a rate constant for removed zombies returning to the zombie state. A larger value of ζ applies when removed individuals are more likely to revert to zombie status.

Here we consider a zombie population attacking a mid-sized town, such as Falkirk in Scotland. As in section 2, we generate artificial data from the ODE model and investigate the quality of the inference. We infer the initial conditions for the zombie and human species, as well as all parameter values. We assume that there are no removed individuals initially, around 40,000 humans living in the town and around 10,000 zombies attacking from Paisley. Given daily observations over a period of $[3, 5, 7, 9]$ days, we consider the *predictive model output*, that is, two standard errors or a 95% confidence for the output of the ODE at each time point, over the 10 days after the first zombie attack, shown in Figure 8. The solid line shows the model output given the posterior mean of the inferred parameters and initial conditions. Figure 9 shows also the inferred initial conditions, which in this example are relatively insensitive to the number of observed data points. As the the number

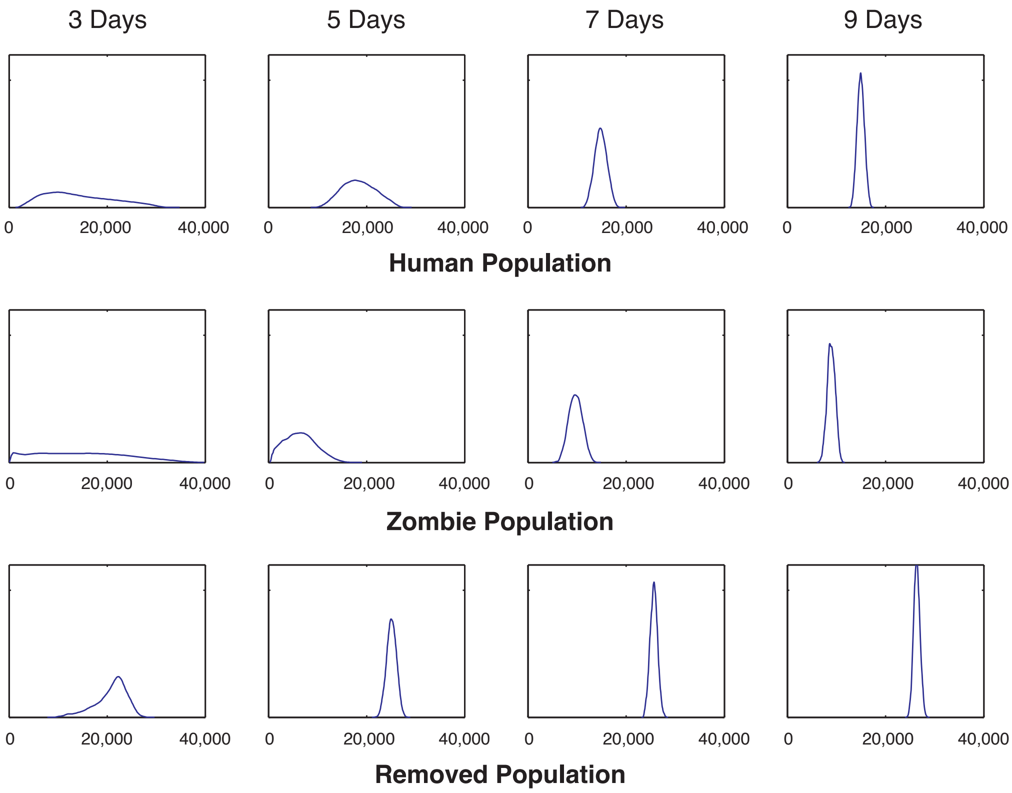

Figure 10: Predicted levels of humans, zombies and removed on day 15 having observed levels of humans and zombies for $[3, 5, 7, 9]$ days.

of observed data points increases, however, we see that the uncertainty in the predictive model output over subsequent days does indeed decrease. This is also seen clearly when we consider the predictive posterior model output for day 15, given observations over $[3, 5, 7, 9]$ days, as shown in Figure 10. With data for just three days we learn very little about the population sizes on day 15 after the zombie attack. Given an additional observation on each of the following two days however, we can predict with much greater certainty that the number of surviving humans is likely to be between 10,000 and 25,000. As we collect more data over the subsequent days, our predictions become more confident, as seen from the more sharply peaked posterior distributions over the each population. Indeed, after 7 and 9 days, the predicted number of humans is roughly between 10,000 and 18,000, and 13,000 and 17,000 respectively. These predicted ranges are tending towards the “true” number of humans, determined by the system of ODEs to be 14,790. Likewise the predicted numbers of zombies and removed individuals tend towards their "true" values of 10,426 and 24,784, respectively.

4 Model Selection

We now consider a second level of inference, in which there is uncertainty not only in the parameters, but also in the specified model. Suppose during the zombie attack on Falkirk that there are rumours that the zombies will only attack in pairs. To simplify things, we suppose that, if the rumours are true, a zombie will always try to flee if it meets a human alone, and, in the same vein, a human who meets a pair of zombies will always try to flee. So

- when a single zombie encounters a susceptible, either both remain unscathed or the zombie becomes removed,
- when a pair of zombies encounter a susceptible, either all remain unscathed or the susceptible succumbs to zombification.

This sounds like a plausible model for behaviour and applying similar arguments to those used for (6)–(8) we may arrive at the alternative set of ODEs

$$\begin{aligned} S' &= -\beta S Z^2, & (9)\\ Z' &= \beta S Z^2 + \zeta R - \alpha S Z, & (10)\\ R' &= \alpha S Z - \zeta R. & (11) \end{aligned}$$

We refer to (6)–(8) as Model 1, and (9)–(11) as Model 2.

Given some data, we may now perform parameter inference over each model. Observations over nine days were generated by simulating from Model 1, in which zombies attack individually, and adding some Gaussian distributed noise with standard deviation 500. We wish to determine whether an inference algorithm will allow us to conclude that Model 1 describes this data better than Model 2, if we did not have any information about the rate constants. Figures 11, 12 and 13 show the posterior model outputs for the two proposed models, setting the standard deviation of the noise to be $[500, 1000, 2000]$ respectively. Visually trying to assess which is the better model is difficult, since the posterior output covers most of the data points for both models.

We may therefore resort to calculating *Bayes factors* such that B_{12} represents the weight of statistical evidence in favour of Model 1 over Model 2. This is computed as the ratio of the marginal likelihoods for the two competing models,

$$B_{12} = \frac{P(\mathbf{Y}|M_1)}{P(\mathbf{Y}|M_2)}. \tag{12}$$

We recall that calculating the marginal likelihood involves estimating the integral of the likelihood times the prior over all values of the parameters, which is an extremely challenging task. We employ the technique of thermodynamic integration, which has recently been shown to provide accurate, low variance estimates of this quantity [4], whereas other seemingly simpler methods, such as the Posterior Harmonic Mean estimator, may fail to produce usable results [2].

Table 1 is a useful guide for interpreting the evidence provided by the estimated Bayes factors [8]. Table 2 shows the results of the marginal log-likelihoods estimated 10 times for each model. In each case, the log Bayes factors correctly identify that Model 1 was used to produce the data. As the standard deviation of the added noise increases from 500 to 1000 to 2000, the weight of evidence as indicated by the log Bayes factors decreases, although the evidence remains substantially in favour of the correct model.

Table 1: Interpretation of Bayes Factor

B_{12}	Evidence against alternative
1 to 3	Not worth more than a bare mention
3 to 10	Substantial
10 to 100	Strong
>100	Decisive

5 Is it Safe to Go Out Yet?

Having introduced the Bayesian statistical machinery for performing inference over systems of differential equations to describe zombie attacks, we now return to a scenario posed in the introduction.

Table 2: Summary of marginal likelihoods for each model

Model	Noise SD	Marginal Log-Likelihood (± Standard Error)	Log Bayes Factor $\text{Log}(B_{12})$
Model 1	500	-152.7 (± 0.1)	31.5 (± 1.8)
Model 2	500	-184.2 (± 1.7)	
Model 1	1000	-158.5 (± 0.1)	16.9 (± 1.1)
Model 2	1000	-175.4 (± 1.0)	
Model 1	2000	-167.1 (± 0.1)	9.9 (± 3.5)
Model 2	2000	-177.0 (± 3.4)	

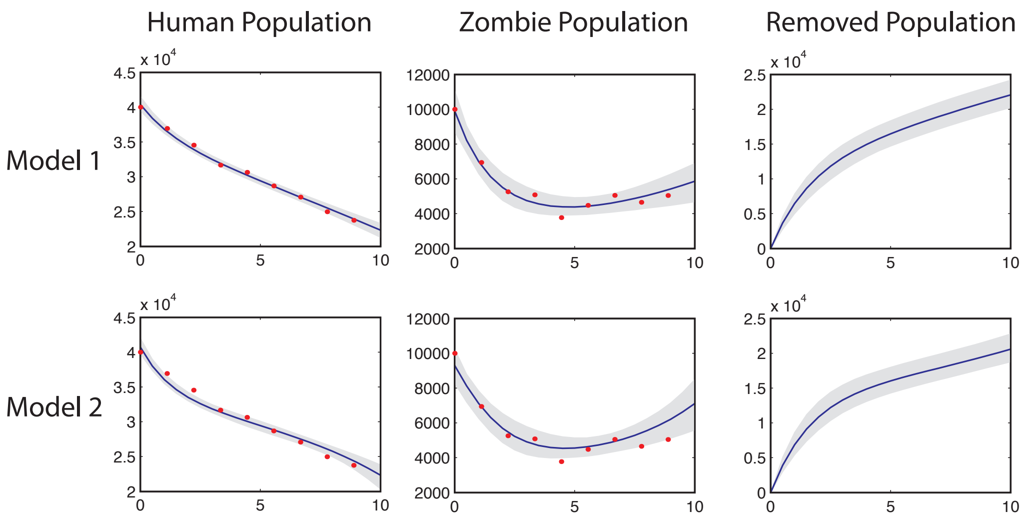

Figure 11: Posterior output for the two competing model hypotheses with the standard deviation of noise set to 500.

- Soldiers patrolling the area have reported daily observed zombie numbers for the past 5 days as $123, 127, 104, 92, 74$. Is it safe to go out yet?

Suppose that we are holed out in the basement of a local convenience store with access to enough Irn Bru and Tunnock's Caramel Wafers to survive for around 50 days. Knowing those zombie levels over the past 5 days, should we hold tight or chance an escape?

We now know how to address this question through the use of Bayesian modelling. We shall assume Model 1 to be a fair representation of the interaction between the zombie population, human population and the removed population. (Of course, if we had multiple plausible models we could once again do full model comparison by calculating Bayes factors.) We perform parameter inference over the model given our data of daily observations of zombies and plot the predictive model output, shown on the left of Figure 14.

In this case, the 95% confidence output from the model includes a wide variety of outcomes, and the uncertainty in the estimate naturally increases with time. We cannot rule out the scenarios where (a) there is relatively little impact on the the human population, or (b) zombies take over completely within the next month! The mean number of zombies continues to decrease over the next 5 days, until day 10 when it begins to increase. We may therefore argue in favor of making an early exit, after which there is much more

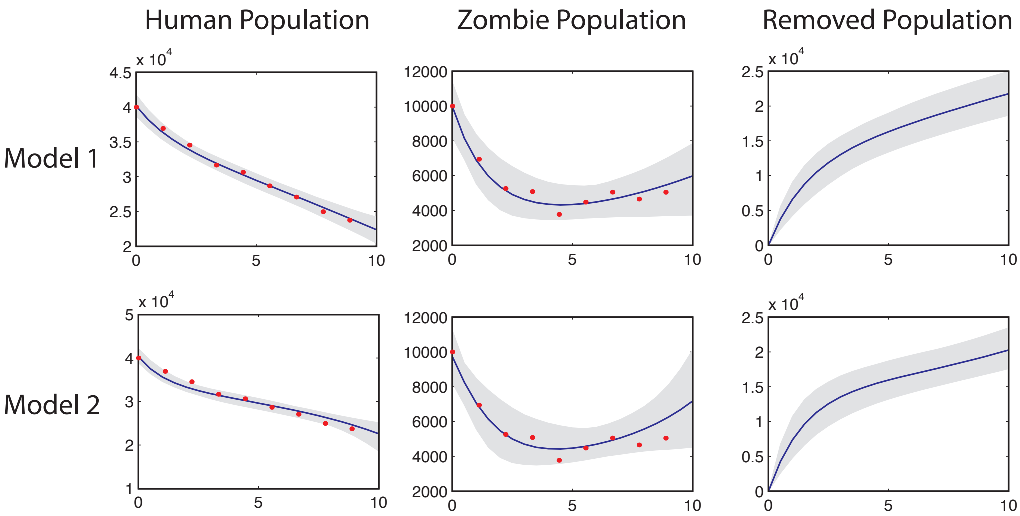

Figure 12: Posterior output for the two competing model hypotheses with the standard deviation of noise set to 1000.

uncertainty that the number of zombies in the village will still be relatively low.

If we have additional data on the human population in the village over the past 5 days, then we may again perform parameter inference over our model including the data for both humans and zombies. The additional data and predictive model output is shown on the right of Figure 14. Given this extra data, we can say with much more certainty that the number of zombies in the village will remain low for a longer period of time, making it perhaps less urgent for us to make our getaway immediately. In this case it is feasible to sit tight a little longer and enjoy our surroundings before gathering up supplies and attempting an escape.

6 Discussion

Mathematical modelling of natural phenomena has a long and illustrious history. Differential equations have the potential to describe and predict the behaviour of many physical and engineered systems. However, any mathematical model represents an abstracted summary that cannot be claimed to capture all characteristics of interest. A well-known quote that is paraphrased from [1] and attributed to George Box says that " All models are wrong but some are useful." Modelling involves compromises and it generates an inherent level of uncertainty. Moreover, the task of identifying

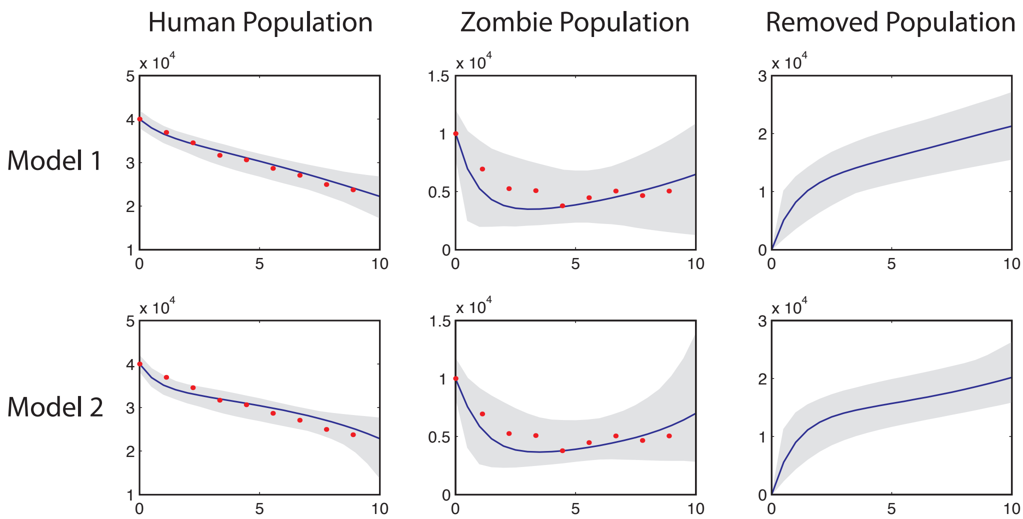

Figure 13: Posterior output for the two competing model hypotheses with the standard deviation of noise set to 2000.

unknown or unmeasured model parameters introduces further uncertainty. In cases where model predictions may be used to guide policy, for example, in economics, weather forecasting and epidemiology, a systematic and consistent treatment of all levels of model uncertainty is vital. We have demonstrated through simple examples in an eye-catching context that the Bayesian statistical framework provides an appropriate means with which to capture information and reason under uncertainty. We have shown how Bayesian inferential methodology may be extremely useful not only when calibrating statistical models based on systems of differential equations, but also in comparing different models and judging which best explains the observed data.

An outstanding challenge in this area is the calibration of models with a large number of unknown parameters, which requires efficient sampling in very high dimensions. Added complications arise when parameters are highly correlated. The underlying 'forward' problem of solving initial-value ODEs numerically is a well-studied topic which has spawned very sophisticated software tools. However, in the inference context it is particularly important to simulate the ODE quickly, since massive numbers of solves may be required, and there is potential for exploiting the special nature of the problem: (a) highly accurate solutions are not required and (b) the ODE may be solved repeatedly for very similar parameter values. Another key issue is the development and analysis of new Markov chain Monte Carlo methodologies to provide unbiased and low variance estimates of the quantities required for

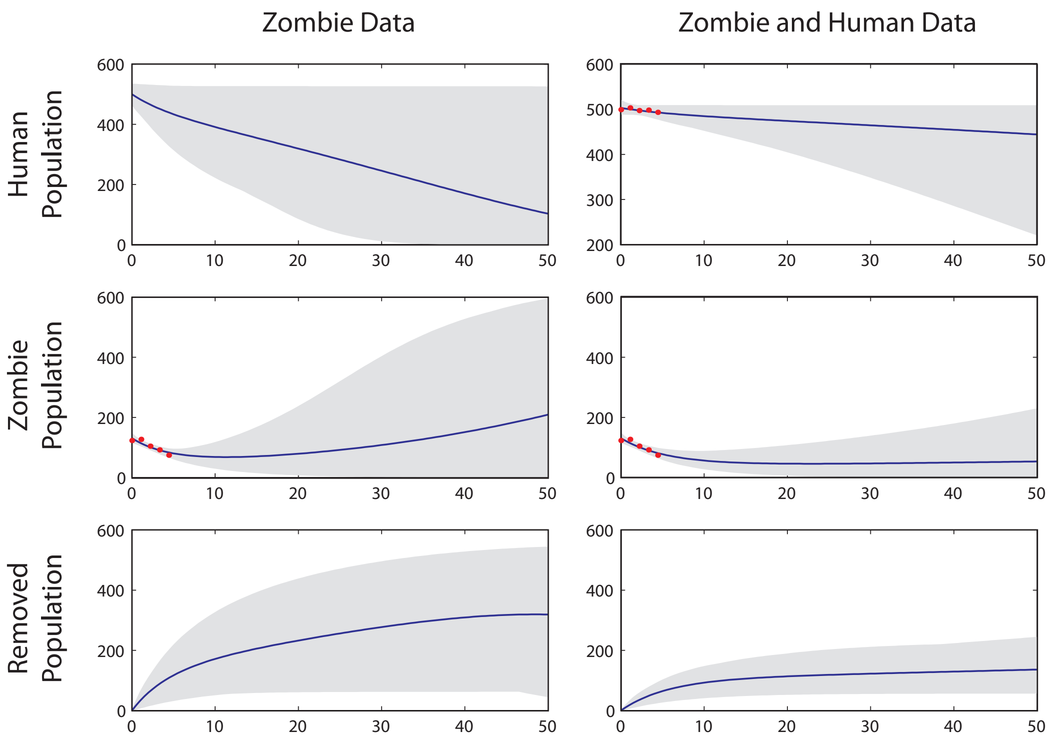

Figure 14: The left hand side plots show the predictive model output from the 3 parameter model given observations of zombies attacking a small village. The right hand side plots show the predictive model output given additional observations of human population numbers.

model comparison.

Overall, this blossoming field brings together many ideas from applied mathematics, statistics and computer science and it offers many opportunities for interdisciplinary research.

References

[1] G. Box and N. Draper, *Empirical Model-Building and Response Surfaces*, Wiley, 1987.

[2] B. Calderhead and M. Girolami, *Estimating Bayes factors via thermodynamic integration and population MCMC*, Computational Statistics and Data Analysis, 53 (2009), pp. 4028–4045.

[3] B. Calderhead, M. Girolami, and N. Lawrence, *Accelerating Bayesian inference over nonlinear differential equations with Gaussian processes*, Advances in Neural Information Processing Systems, 21 (2009), pp. 217–224.

[4] N. Friel and A. Pettitt, *Marginal likelihood estimation via power posteriors*, Journal of the Royal Statistical Society: Series B, 70 (2008), pp. 589–607.

[5] W. Hastings, *Monte Carlo sampling methods using Markov chains and their applications*, Biometrika, 57 (1970), pp. 97–109.

[6] D. Hitchcock, *A history of the Metropolis-Hastings algorithm*, The American Statistician, 57 (2003).

[7] E. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University Press, 2003.

[8] R. Kass and A. Raftery, *Bayes factors*, American Statistical Association, 90 (1995), pp. 175–196.

[9] P. Munz, I. Hudea, J. Imad, and R. Smith?, *When zombies attack!: Mathematical modelling of an outbreak of zombie infection*, in Infectious Disease Modelling Research Progress, J. Tchuenche and C. Chiyaka, eds., 2009, pp. 133–150.

[10] D. Sivia and J. Skilling, *Data Analysis: A Bayesian Tutorial*, Oxford University Press, 2006.

[11] V. Vyshemirsky and M. Girolami, *Bayesian ranking of biochemical system models*, Bioinformatics, 24 (2008), pp. 833–839.

[12] ——, *BioBayes: A software package for Bayesian inference in systems biology*, Bioinformatics, 24 (2008), pp. 1933–1934.