Todos

University of Auckland

Literature Review

Parametrisation and Identification Methods for Epidemic Systems

David Wu

supervised by

Dr. Oliver Maclaren

and

Dr. Vinod Suresh

Department of Engineering Science

November 2019

Contents

List of Figures

1 Overview

Epidemiological modelling arose from the early attempts at analysing the spread of malaria by Ross [1] and empirical modelling by Kermack and McKendrick [2] in the early 20th century. These models were simple, often analytically tractable, and hugely influential in the prediction and control of infectious diseases. In fact, the SIR models introduced by Kermack and McKendrick are still used in the analysis of infectious outbreaks to this day.

One question that has been posed recently is how well predictions can be made from these models. Studies in forecasting proactively , and more commonly retroactively, [NeedRef]have been performed for epidemics of, for example, HIV, SARS, influenza, Zika and Ebola [NeedRef]. Other, more computationally intensive, statistical approaches have also been taken to explain the spread of the same diseases. This has led to the creation of monitoring and prediction tools such as FluSight [NeedRef]and Google Flu Trends [NeedRef], to varying degrees of success. **[INCOMPLETE]**

The major challenges in the field stem from the mathematical problem that is being solved. At a high level, the problem can be constructed as an objective function, where the model is matched to data . The challenges with this objective function are two-fold:

- the objective function is highly non-convex and discontinuous, and
- the problem is likely (practically) non-identifiable.

Additionally, the need to quantify the uncertainty in the predictions further complicates the problem. Methods derived from frequentist frameworks can be used to construct confidence intervals based on local approximations of the objective, but often this approximation can break down for the limited datasets we obtain. Conversely, Bayesian methods can be used to provide more robust estimates, but their need to perform many forward model simulations mean that is is extremely computationally expensive. Both approaches also have common inherent assumptions about the model which can cause estimability problems. **[INCOMPLETE]**

This motivates the use a variety of methods, which we will touch on in this review, to combat different aspects of this problem. **[INCOMPLETE]**

Review Outline

This literature review will focus on the models in mathematical epidemiology, and methods for parametrising these models. The review is structured as follows. In Section 2, we will examine literature covering the temporal evolution of infectious diseases (Section 2.1) and their extension into further (spatial) dimensions (Section 2.2). Next, we examine methods of parametrisation for the general class of models discussed in Section 3. Section 4 will cover a sample of applications of these methods. Finally, we discuss some gaps in the literature in Section 5, and summarise the findings in Section 6

2 Epidemiological Models

Epidemiological models are varied in form and motivation. Some formalise the causal relationships between different effectors in a system [NeedRef]. These are useful for identifying potential avenues for intervention ——————[INCOMPLETE]—————— . Of more interest to this review is the class of models that describe the evolution of the epidemic in time (and potentially other dimensions). Of this, there exist two subclasses of models - statistical and compartmental. Statistical models typically generate interpolating regressions on the data; compartmental models are generally derived mechanistically. ——[INCOMPLETE]——

2.1 Temporal Models

Much of the analysis of epidemiological models is done for ordinary differential equations (ODEs). These are models of the form

$$u_t = f(u;\theta). \tag{2.1}$$

For example, the (non-age-dependent) classic SIR model [2] can be represented this way:

$$\begin{aligned} u &= (S,I,R)^T, \\ \theta &= (\alpha,\beta)^T, \\ f(u;\theta) &= \begin{pmatrix} -\beta SI/(S+I+R) \\ \beta SI/(S+I+R) - \alpha I \\ \alpha I \end{pmatrix}. \end{aligned} \tag{2.2}$$

where the states are **S**usceptible, **I**nfected/Infectious, and **R**emoved/Recovered. The parameters α and β represent the rate of recovery and effective transmission rate respectively; typically $\lambda \coloneqq \beta I/(S+I+R)$ is termed the force of infection, and $N \coloneqq S+I+R$ is the total population size.

Most compartmental models are extensions on this representation, introducing additional states of u, or changing the form of f. Much analysis has been performed for this class of models; we will cover some basic concepts.

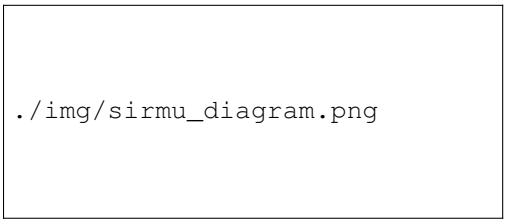
One of the most important quantities in this class of model is the basic reproduction number R_0. This represents the number of secondary infections caused by a single infected individual, and acts as a threshold value for the model. If this value is above 1 then an outbreak will occur, where a large number of individuals will become infected. For the simple SIR model, the associated R_0 would be $\frac{\beta}{\alpha}$, i.e. the ratio between rate of effective contacts and rate of recovery. The interpretation is that is if the infected manages to recover before being able to cause a secondary infection, then the disease will naturally die out.

The dynamics become more complicated with additions of elements that replenish the susceptible pool. A typical modification is the addition of vital dynamics [3], which replenishes the

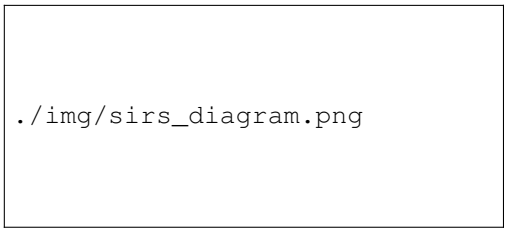
susceptible pool with births. Making a constant population assumption, we can write this as:

$$\begin{aligned} \dot{S} &= -\beta SI/N - \mu S + \mu N, \\ \dot{I} &= \beta SI/N - \alpha I - \mu I, \\ \dot{R} &= \alpha I - \mu R, \\ N &= S + I + R, \end{aligned} \tag{2.3}$$

for which the quantity $R_0 = \frac{\beta N}{\mu+\alpha}$ can be computed. This form allows for the epidemic to persist endemically ($t \to \infty, I \not\to 0$), as the susceptible pool grows like μR when the infected pool approaches depletion, allowing for a recurrence of infections. A similar effect is seen if waning immunity is assumed, where recovered individuals return to the susceptible pool, either directly or via an intermediary partially susceptible state [4]. These two variants are presented diagrammatically in Figure 2.1.



(a) SIR with vital dynamics



(b) SIRS model (waning immunity)

Figure 2.1: Diagrammatic representations of select variants of the SIR model

This endemic presence of infections and recurrence cycle can also be exacerbated by resonance, when the parameters are nonstationary [5]. A sinusoidal form of the transmission parameter $\beta = \beta_0 + \beta_1 \cos(2\pi t)$ in a waning immunity model can induce resonance for particular transmission parameters, causing large oscillations in the number of infected over time. This can interpreted as periodic epidemic outbreaks, which is what is observed for endemic epidemics such as influenza. However, this introduces more discontinuities in parameter space, as these resonance frequencies can represent suddenly well-fitting, or misfitting, trajectories. This idea of seasonal forcing is further extended by Shaman and Kohn [6] for influenza dynamics, by considering humidity as a possible driving force behind the seasonal transmission. This motivates the simulation of such recurrent epidemics for a changing climate, such as the study done in [7].

Another idea in epidemiology is modelling structures in the population. A common form is the age-structured model [8], [9], where the state variables are also partitioned by age groups. Models are typically of the form

$$\begin{aligned} \dot{S}_i &= -\sum_j \beta_{ij} S_i I_j / N, \\ \dot{I}_i &= \sum_j \beta_{ij} S_i I_j / N - \alpha I_i. \end{aligned} \tag{2.4}$$

This is useful in examining heterogeneities in the transmission properties within and between different age groups. This allows for a coarse-grain modelling of the differences in contact networks in each of the groups. They can also represent differences in the susceptibility of a group, perhaps due to lower rates of vaccination. A similar spatial idea will be visited in Section 2.2.2, with spatial mixing heterogenities.

With such increasing complexity of epidemiological models, the recovery of the basic reproduction number, R_0, becomes more complicated. Consider the general model with m non-disease compartments and n disease compartments. Let $x \in \mathbb{R}^n$ be the disease compartments, and $y \in \mathbb{R}^m$ be the non-disease compartments, with

$$\frac{dx_i}{dt} = \mathcal{F}_i(x, y) - \mathcal{V}_i(x, y),\ i = 1, \ldots, n\,. \tag{2.5}$$

Let $(0, y_0)$ represent the disease-free equilibrium ($y \to y_0$ as $t \to \infty$). It can be shown that R_0 can be recovered [10] as the dominant eigenvalue of the *next generation matrix*

$$K_L = \frac{\partial \mathcal{F}_i}{\partial x_j}(0, y_0) \left(\frac{\partial \mathcal{V}_i}{\partial x_j}(0, y_0) \right)^{-1}.$$

In addition to determining the threshold relations, the eigenvector associated with the dominant eigenvalue can also be used to approximate the dominant modes of infection in an outbreak. Other model structures, such as cross-infection [11], [12] or heterogeneous mixing [13], make use of this approach in order to derive thresholds. Some work has been done on direct parametrisation on the next-generation matrix, in lieu of the rate parameters in the model [9].

The above is not an exhaustive summary of the different temporal compartmental models and related concepts in mathematical epidemiology (we have merely covered some simple examples for a few scenarios). Other variants, such as the SEIR, SITR

There also exist stochastic variants of the *deterministic* models presented above. These vary in complexity, but a simple stochastic variation is typically to consider each transition from one state to another as an event, that occurs with some probability derived from its rate [14]. Then, events are drawn and simulated consecutively with the Gillespie method or similar to generate the trajectories in time as a Markov chain. Although these models can asymptotically converge towards the trajectories of their deterministic counterparts, their stochastic nature can lead to additional behaviour, such as extinction, that is not predicted by deterministic models [15].

2.2 Spatio-temporal Models

Though purely temporal models are common in epidemiology, spatio-temporal models also have a rich history. Perhaps the first rough survey of this subject was performed by John Snow [16] on the spread of cholera in London. Spatial proliferation of epidemics is of particular interest when considering the containment and control of an emerging epidemic. Studies range in scale from local city-to-city models [17] to international models[18], [19], and remain less explored than their non-spatial counterparts. More interest has been placed in spatial models as the changes in human movement and distribution are impacting the way that diseases can spread[20].

2.2.1 PDE Models

Perhaps the most natural extension to the compartmental models above is to represent spatial models as partial differential equations (PDEs). The earliest forms of these models was perhaps introduced by Kendall [21]. In this paper, a spatial model is derived from an averaging of the effect of nearby infected on the transmission of the infection to the susceptibles. This gives rise to the following model:

$$
\begin{aligned}
S_t &= -\beta S\left(I + kI_{xx}\right), \\
I_t &= \beta S\left(I + kI_{xx}\right) - \alpha I, \\
R_t &= \alpha I.
\end{aligned} \tag{2.6}
$$

Travelling wave solutions of this PDE were then sought, and shown to be reducible to a differential equation of I with respect to S. It is then shown that travelling wave solutions only exist when the wave velocity is above a threshold wave velocity $2\sqrt{\beta k(\beta - \alpha)}$, using arguments from the solution of the Fisher-Kolmogoroff equation [22]. A similar derivation is made in Murray [23] with a subtly different system applied to rabies, where the spatial derivative is motivated simply as a diffusion process of the infected. A thorough analysis of restricted forms of partial differential equation models is presented in Rass and Radcliffe [24].

2.2.2 Metapopulation Models

A different way of conceptualising spatial models is the metapopulation models. This is done by discretising the spatial domain into different spatial compartments, and modelling the movement of people between these compartments. This means that the movement of individuals do not need to be defined over the entire domain, which is often difficult, due to the nature of human migration[NeedRef]. Further, it reduces the entire spatial domain into natural discretisations, meaning standard methods of dealing with ODE compartmental models can be used instead. A typical metapopoulation model is given by Hyman and Laforce [25] for inter-city spread of influenza:

$$
\frac{dS_i}{dt} = -\beta_S S_i \frac{I_i}{N_i} + \eta R_i + \mu(S_i^0 - S_i) + \sum_j \left(m_{ji}\frac{S_j}{N_j} - m_{ij}\frac{S_i}{N_i} \right), \tag{2.7a}
$$

$$
\frac{dI_i}{dt} = \beta_S S_i \frac{I_i}{N_i} - \alpha I_i - \mu I_i + \sum_j \left(m_{ji}\frac{I_j}{N_j} - m_{ij}\frac{I_i}{N_i} \right), \tag{2.7b}
$$

$$
\frac{dR_i}{dt} = \alpha I_i - \eta R_i - \mu R_i + \sum_j \left(m_{ji}\frac{R_j}{N_j} - m_{ij}\frac{R_i}{N_i} \right). \tag{2.7c}
$$

This includes vital dynamics with a carrying capacity in each city i of S_i^0, waning immunity at a rate of η, and explicit modelling of migration between cities at rates m_{ij} from city i to city j. The model assumes that the time scale is resolved enough so that infections are between individuals in the same city, as indicated by the form of the force of infection. This leads to a large $4 \times i$-dimensional system that can be simulated. Further, $\mathcal{R}_0$ can be computed for each of the patches, if say $\beta_{\{S,P\}}$ were different in each patch. Multi-strain variants of this model are presented in [17], [26].

An alternative formulation is suggested by Bichara *et al.* [27][28], where the explicit migration terms are removed in favour of a residence-vistor setup:

$$
\frac{dS_i}{dt} = \mu(S_i^0 - S_i) + \eta R_i - \sum_j \left(\beta_i p_{ij} S_i \sum_k p_{kj}\frac{I_k}{N_k} \right), \tag{2.8a}
$$

$$
\frac{dI_i}{dt} = -\alpha I_i - \mu I_i + \sum_j \left(\beta_i p_{ij} S_i \sum_k p_{kj}\frac{I_k}{N_k} \right), \tag{2.8b}
$$

$$
\frac{dR_i}{dt} = \alpha I_i - \mu R_i - \eta R_i, \tag{2.8c}
$$

where p_{ij} denotes the probability that a resident of city i is in city j. Here, the force of infection is also 'local', but x_i now tracks the states of the *residents* of city i, and the emphasis of movement is on visitations, as opposed to permanent relocation, as above. It also means that an individual's probability of migration is tied to their city of residence, as opposed to location at the time of movement.

One more common idea in metapopulations models is the gravity model of movement. This is an encapsulation of the visitation models above, but with an explicit form of p_{ij}. For example, [29] uses the equivalent of:

$$p_{ij} = \kappa_{ij} \frac{N_i N_j}{{d_{ij}}^2}, \tag{2.9}$$

where d_{ij} is the Euclidean distance between the cities i and j, and κ_{ij} is a weighting term. This model has been applied in a New-Zealand specific context in the sociology literature [30], using data generated from Census information. Particular attention was paid to the rural-urban and international migration patterns. Some epidemiological literature, however, discusses missing gaps in the gravity model. Truscott and Ferguson [31] shows that infrequent, long-distance visitations are not captured, as well as movements involving nodes with small populations. Li *et al.* [19] notes that the model is valid only if the duration of the epidemic is significant; otherwise, local effects can dominate. This is congruent with observations of early-epidemic behaviour, and sparse contact networks.

2.2.3 Individual-based Models

To capture the stochasticity, sparsity and heterogeneity of contact networks, stochastic variants of the deterministic models can be used. A more common approach is to model individuals in the population separately. These models allow for rich contact frameworks to be modelled. This removes the assumptions of well-mixing from the model, and can be used to explain the non-exponential growth of the epidemic curve in early time. Within this class, there are also multiple levels of generalisation. Network models, for example, assume homogenous contact over all connections between individuals. This allows for contact between individuals to be simulated simultaneously. This homogeneity is then weakened in individual-based models (also agent-based models). Of course, the increase in modelling complexity and richness is offset by computational expense. Cost-saving methods, such as the Gillespie method that can be used for stochastic variants of deterministic models, are now much more difficult to implement correctly, given the heterogeneity of event occurrences.

2.2.4 Challenges

The central problem for spatial epidemic models is that the modelling of (heterogeneous) contacts is extremely hard.

2.3 Other Models

2.3.1 Statistical Models

Another class of model is most commonly seen in statistical literature for prediction of epidemics. Statistical regression models can be constructed with no direct modelling of the underlying dynamics, and used to predict incidence of recurring epidemics. For example, [NeedRef - flu]

uses a linear regression model to predict the incidence of influenza across the continental United States. [INCOMPLETE]

2.3.2 Phenomenological Models

There are also phenomenological models that describe the incidence of disease, similar to statistical regression models. The most popular one used for early-epidemic behaviour [32], [33] is:

$$\dot{C} = rC^p, \; p \in [0, 1]. \tag{2.10}$$

Here, C represents the count of incidences. The bounds on p mean that the model exhibits sub-exponential ($p < 1$) or exponential growth ($p = 1$). Because early growth of epidemics is sensitive to the contact network of the initially infected, polynomial growth is often observed, for example in the 2014 West African cases of Ebola [33].

3 Parametrisation Methods

We now turn our focus to methods for parametrising these models with respect to data. The problem we are typically trying to solve is determining the the set of parameters θ for a model $\mathcal{M} : \mathcal{D}u - f(u;\theta) = 0$ such that the model outputs ——[INCOMPLETE]—— i.e. we are interested in the inverse problem to the forward problem of examining the behaviours of the models of the previous section. Bit of a misnomer - the analysis of the models is attempting to determine general forms for when stuff is important, whereas parametrisation is an attempt to [INCOMPLETE] ——————————————————————

3.1 Optimisation-Based Approaches

Optimisation approaches are a natural way of constructing ——[INCOMPLETE]——

3.1.1 Trajectory Matching

The typical approach is nonlinear least-squares. That is, to minimise the sum of squared errors between the model output and data. For an ODE model, this would be minimising the error between the integral (trajectory) of the model in time and the data. We will term this trajectory matching, following Ramsay and Hooker [34]. It can be expressed as finding the optimal parameters θ_{opt} such that

$$\begin{aligned} \theta_{opt} &= \arg\min_{\theta} \sum_i \sum_n \left(y_i(t_n) - S_i(t_n)\right), \\ S(t) &= \int_{t_0}^{t} f(S(\tau);\theta) d\tau, \\ S(t=t_0) &= x_0. \end{aligned} \tag{3.1}$$

where i indexes the state variables (elements of x).

One problem with trajectory matching is that the objective function is often non-convex. Further, we expect that discontinuities exist, corresponding to bifurcations in the model. This leads to a objective that is ill-suited to standard gradient-based (Newton-Raphson) methods or derivative-free methods. Despite this, there exist many common methods that attempt to tackle these problems with the aid of computational brute force. One example is the use of multi-start methods which perform multiple runs of the optimisation procedure from different initial iterates [29], [35], [36]. Methods such as the Latin hypercube sampling technique are used to generate these initial iterates. The hope is that, with sufficient coverage of the parameter space, the minima computed over all the samples will be close to the global optimum. Novel techniques have been proposed [37] to generate initial iterates that are suitable to mitigate this issue.

A secondary problem with the trajectory matching problem is that there is significant computational effort involved in calculating the model outputs, as this typically involves the numerical integration of the model. Due to problems such as stiffness, or complex time-dependent model forms, this can take an unacceptably long amount of time. Hooker *et al.* [38] notes that in their study (described below in Section 4) that a sampling based approach would have been infeasible due to the time it would take to generate a single proposal of the state trajectories. Further,

careful selection of the initial iterates, and subsequent estimates, is required to minimise the chance that the parameters estimated require a long time to integrate.

One philosophical problem with the trajectory matching method, however, is the assumption made about the

3.1.2 Gradient Matching and Splines

An alternative approach to trajectory matching is sometimes proposed to mitigate the issues of this latter problem. Instead of evaluating the model outputs via integration, the associated derivatives of the data can be taken, and compared to the model directly. This method, termed by Ramsay and Hooker [34] as gradient matching, can be expressed:

$$\theta_{opt} = \arg\min_{\theta} \sum_{i} \sum_{n} \left(\mathcal{D}\hat{y}_i(t_n) - f(\hat{y}_i(t_n);\theta)\right)^2, \tag{3.2}$$

for some model

$$Dx = f(x;\theta),$$

where $\hat{y}$ represents an approximation to the data y. This approximation $\hat{y}$ is typically a smoothed representation of the data, and is typically expressed as a basis expansion

$$\hat{y}(t) = \sum_{k} C_k \Phi_k(t), \tag{3.3}$$

in some basis Φ. This can then be substituted in the objective function, giving the second stage:

$$\theta_{opt} = \arg\min_{\theta} \sum_{i} \sum_{n} \left(C\mathcal{D}\Phi(t_n) - f(C\Phi(t_n);\theta)\right)^2. \tag{3.4}$$

Φ is often a spline basis, and the coefficient vector C is computed from minimising the least squares error

$$\|y - C\Phi\|^2. \tag{3.5}$$

However, there exists a meta-problem with the use of splines, in the fact that the quality of the fit (when done with least-squares) is quite sensitive to the choice in number and location of the knots of the spline. In short, a low number of knots will typically lead to underfitting, and a high number of knots leads to overfitting. A now standard solution to this problem is to penalise the second derivative of the fit to enforce smoothness. This approach, presented in such texts as [39]–[41], modifies Equation 3.5 to

$$\|y - C\Phi\|^2 + \lambda \|C\Phi_{tt}\|^2, \tag{3.6}$$

with some smoothing parameter λ. Wahba [39] shows that a suitable value of λ can be computed by minimising the generalised cross validation error (weighted leave-one-out error)

$$V(\lambda) = \frac{1}{n}\sum_{k=1}^{n}(y_k - C\Phi)\textbf{INCOMPLETE} \tag{3.7}$$

One of the problems with using gradient matching in a practical context is that the seconds step, computing $C\mathcal{D}\Phi(t)$, requires a fully observed system. That is, if the model has m state variables, all m states must be observed at each time point. When considering with real data, it is often impractical, if not impossible, to gather data on all states. There have been attempts at estimating these unobserved states in the econometrics literature [NeedRef]through the instrument variables method.

3.1.3 Generalised Profiling

In [42], and expanded in [34], Ramsay et al. introduce the concept of generalised profiling, which attempts to solve the (complementary) issues of both trajectory matching and gradient matching. The crux is that the optimal parameter(s) θ_{opt} is the minimiser of

$$\|y - \Phi c\|^2 + \lambda \|\mathcal{D}\Phi c - f(\Phi c; \theta)\|^2 . \tag{3.8}$$

This is extremely similar to the smoothed spline function (Equation 3.6), replacing the smoothing term with the gradient matching term. This is because the motivation is very similar. The smoothing term penalises non-smoothness (i.e. non-zero second derivatives), which generally does not hold for most models. Intuitively, we can simply replace this with our model, and penalise for that. From a different viewpoint, this can also be constructed as a relaxation of the constrained trajectory matching problem:

$$\begin{aligned} &\min \|y - \Phi c\|^2 , \\ &s/t \\ &\quad \|\mathcal{D}\Phi c - f(\Phi c; \theta)\|^2 = 0. \end{aligned} \tag{3.9}$$

In [34] (and all related works [38], [42]–[44]) the parameter estimation is split into two stages, with different objective functions. The above objective function is used in the inner optimisation procedure, where the spline coefficients c are estimated for a given θ; the standard least-squares objective is used to in an outer optimisation procedure where the optimal spline coefficients are used as a surrogate for the estimated state. We can express this as:

$$\theta_{opt} = \arg\min_{\theta} \left\{ \|y - \Phi c(\theta)\|^2 \right\} \qquad =: \arg\min_{\theta} H(\theta), \tag{3.10}$$

$$c(\theta) = \arg\min_{c} \left\{ \|y - \Phi c\|^2 + \lambda \|\mathcal{D}\Phi c - f(\Phi c; \theta)\|^2 \right\} \qquad =: \arg\min_{c} J(\theta, c). \tag{3.11}$$

It can be shown, using the implicit function theorem, that the derivative of the outer objective $\frac{dH}{d\theta}$ can be expressed as a function of known values, such as the partial derivatives of J. This then allows derivative methods for minimisation to be used. Further, since $c(\theta)$ is likely to be locally smooth and continuous, the inner optimisation procedure should take a minimal amount of iterations to converge.

We observe, however, that the outer objective is strangely formulated. If we consider the inner objective as profiling out the structural parameters, it makes more sense for the outer objective to have the same form as the inner objective. The reasoning given by Ramsay and Hooker [34] is that $c(\theta)$ represents a regularised estimate of the state, so further regularisation in the outer objective is not needed. Perhaps a stronger (but equally non-satisfying) reason would be that the form of the outer objective lends more easily to an interpretation of the objective as the log-likelihood of a simple normally distributed error model, which allows for straightforward construction of confidence intervals for uncertainty quantification (see Section 3.1.5).

3.1.4 Choices of Regularisation Parameter

We see that in Equations (3.6) and (3.8) that an extra tuning parameter (λ) is introduced. This parameter describes the tradeoff between the fit to data term and another regularisation term (smoothing or model fit). It can be shown [NeedRef]that the solution to the objective will be dependent on the choice of this parameter. That raises the question of how to choose the value

of such a parameter. From the inverse problem literature, we see that there are a multitude of methods of choosing this value. They can classified into *a priori* and *a posteriori* methods.

A priori methods define a strict stopping condition based on prior knowledge or estimates of the error structure and size present in the model. One of the most popular techniques in this class is the Morozov discrepancy principle. **[INCOMPLETE]** However, an estimate of the size of the error is generally not available, so such methods cannot be applied.

A posteriori methods subvert the need for a prior estimate of the error by examining the properties of the estimate after it has been computed. Methods of this class include the L curve criterion and (generalised) cross-validation techniques. The L curve criterion is a form of Pareto front of the two objectives of solving the inverse problem, and regularising the solution. The argument is that the regularisation is seen to provide value, until it significantly changes the quality of the fit, thus the "optimal" choice of regularisation parameter will be at this value threshold. This corresponds geometrically to the area where the L curve turns, and consequently to where the curvature is the highest. It is, by all accounts, an ad hoc method for tuning the regularisation parameter. Cross-validation techniques **[INCOMPLETE]**

The problem of selecting these tuning parameters is exacerbated when more regularisation terms are included. For example, if we add a Tikhonov regularisation term to the generalised profiling objective in Equation (3.8), we would also need to add an additional tuning parameter, which would need to be selected alongside λ. Methods have emerged for such multi-parameter selection problems; we focus on L-curve methods for parameter selection. **[INCOMPLETE]**

3.1.5 Uncertainty Quantification

Much of the uncertainty estimates of the optimisation-based methods are generated in a frequentist manner. The observation is that if we specify the likelihood function as

$$\mathcal{L}(y(t)|\theta) \sim \mathcal{N}(S(t), \Sigma), \tag{3.12}$$

where $S(t)$ is defined as in Equation 3.1, and Σ is some covariance matrix; then the negative log-likelihood becomes the sum of squares function for $\Sigma = \sigma^2 I$. That is, minimising the negative log-likelihood will be identical to performing trajectory matching. This is powerful, since it allows standard asymptotic statistical theory to be applied to the estimates produced by trajectory matching. In essence, this means that the uncertainty of the estimate can be quantified.

The standard approach is to compute the Fisher information matrix from the Hessian of the objective function[NeedRef]. This can be then inverted and used to construct confidence intervals of arbitrary confidence thresholds, as specified by the χ^2 distribution[NeedRef]. Of course, these intervals are constructed by locally approximating the objective function (and thus likelihood).

A similar idea is explored in the literature, often termed *sensitivity analysis* [NeedRef]. Such methods determine the forward and adjoint responses of a model to some perturbations in the input or output, respectively. This is subtly different to the

3.2 Bayesian Approaches

3.2.1 Bayesian Methodology

A standard framework for interpreting the uncertainty in parameter estimates is the Bayesian framework, which centres around Bayes' theorem:

$$P(\theta|y) \propto P(y|\theta)P(\theta). \tag{3.13}$$

Here we interpret y as the data, and θ as the parameters we are interested in estimating. In many other inverse problems, the main focus of the problem is the estimation of the state of the system - this will be discussed in Section 3.2.3.

$P(\theta)$ is denoted the *prior distribution* of the parameters, and is a representation of prior knowledge of the parameters. It is also a source of bias in our estimate, so much effort is often made in statistical circles to construct uninformative priors that do not bias the posterior [NeedRef]. $P(\theta|y)$ is the *posterior distribution*, or our estimate distribution of the target parameters conditioned on the knowledge of the data. This makes up our estimate (and representaion of uncertainty) of the model parameters that "fit" the data.

$P(y|\theta)$ is denoted the likelihood distribution. This is a representation of the probability that the data y would be produced given some value of the parameters θ. For model-driven problems, this is often a function of the model outputs. For example, in [45], the likelihood is constructed as a Poisson distribution, centred about the underlying state.

3.2.2 Sampling Methods

The direct computation of the posterior distribution from the likelihood and prior distributions is difficult, often infeasible. In lieu of this, sampling methods aim to generate a set of candidates whose distribution converges on the target posterior distribution.

3.2.3 Data Assimilation Methods

3.3 Estimability Problems

Shitty section, needs massive rewrite

Despite these numerous approaches to parameter estimation, there exist some underlying problems that aren't directly answered. The most significant one is the problem of identifiability. This is where, regardless of the quantity of data provided, the quality of the data or the structure of the model prevents the parameters to be estimated uniquely. That is, there exist at least two distinct set of parameters that can produce indistinguishable observations. [NeedRef - Raue maybe]

We can classify different levels of identifiability. The weakest class, structural identifiability, is the requirement that given an infinite amount of data, that the parameters can be uniquely recovered. Techniques, such as DAISY,

Although the theory works well in principle, there are difficulties that arise when applying it practically.

One underappreciated problem is one of identifiability.

Definition 1 (Identifiability) *A process model, $\mathcal{M}$, under an observation model, $\mathcal{H}$, is said to be* identifiable *if: $\mathcal{H}(\mathcal{M}(\theta_1)) = \mathcal{H}(\mathcal{M}(\theta_2)) \implies \theta_1 = \theta_2$ for all parameter combinations θ_1 and θ_2.*

In particular, we will consider the form

$$\dot{x}(t) = f(x(t), u(t), \theta) \leftrightarrow x(t) = \mathcal{M}(\theta), \tag{3.14}$$
$$y(t) = g(x(t), u(t), \theta) \leftrightarrow y(t) = \mathcal{H}(x(t)). \tag{3.15}$$

The interpretation is that each set of observations has a unique corresponding parameter set. Identifiability that is non-trivial to solve, since it has to do with the structure of the model, or the quality of the data — both difficult things to modify in the parameter estimation process. Perhaps unfortunately, it is easy to show that even simple models can be non-identifiable. The simple turnover model is one such example.

$$\begin{aligned} \dot{x} &= k_1 - k_2 x, \\ y &= sx. \end{aligned} \tag{3.16}$$

Rearranging the equations, we can simplify the model to

$$\dot{y} = k_1 s - k_2 y. \tag{3.17}$$

If we consider that y is the only observable state, then it is obvious that there exist an infinite set of pairs $\{(k_1, s) : k_1 s = c \in \mathbb{R}\}$ for some constant c, and thus an infinite set of parameters that can reproduce the same data. More examples appear in the literature, for a wide range of biological applications [46]–[49] where the system much more complex, and such non-identifiabilities are much more difficult to detect.

Identifiability can be considered as two different problems: structural identifiability — given an infinite amount of data, does the model allow for the parameters to be uniquely recovered, such as the turnover model above?; and practical identifiability — given some finite amount amount of data, can we uniquely recover the parameters?

There exist algebraic methods for analysing the structural identifiability of systems a priori. Differential algebraic techniques such as Differential Algebra Identifiability of Systems (DAISY) or Exact Arithmetic Rank (EAR) methods. The former of these generates a system of algebraic nonlinear equations in θ from the algebraic elimination of the state variables, which can then be analysed for multiple roots (solutions) to detect global non-identifiability. The latter computes the Jacobian of the observation model as a function of the initial state and parameters, and applies the inverse function theorem to use the rank of the Jacobian matrix about a random point to determine local identifiability.

Another method of analysing non-identifiability is the profile likelihood method as introduced by Raue *et al.* [50]. The profile likelihood analyses practical identifiability of systems by computing the change in the objective function of the optima as one constrained parameter is varied, also termed the profile of the likelihood, p:

$$p_i(\theta_i) = \min_{\theta_j, j \neq i} \mathcal{H}(\mathcal{M}(\theta)). \tag{3.18}$$

Intuitively, as θ_i is varied for an *identifiable* model, we expect that the value of the profile will increase, since there is no other parameter that can vary to compensate for the change in θ_i. Consider, however, a non-identifiable model. If there is some non-identifiable combination of parameters, then the change in θ_i can be compensated for by a change in θ_j. Thus, the profile as θ_i will vary negligibly as θ_i is varied.

Bit ugly

This profile is thus the subset of the parameter space that minimises the objective function for

something about the sensitivity analysis approaches. The problem with sensitivity approaches is that they tend to be derived analytically from derivative information. Thus, they perform local approximations to the objective, which can potentially not capture important information. This is equivalent to using the statistical methods to approximate the likelihood function to compute asymptotic confidence intervals.

Identifiability is closely related to the ideas in Section 3.1.5. The ideas of (frequentist) uncertainty quantification are essentially local approximations to the profile likelihood. »Similar ideas in [51], where Monte Carlo methods are used to generate a profile of the parameter estimates as the (known synthetic) data is perturbed with observational noise. This generates correlation plots between parameters, which can be used, similar to the profile likelihood to determine whether or not a parameter is identifiable. Instead of examining the confidence intervals, a derived sample correlation between parameters is used, alongside the computed sample variance, relative to the variance in the measurement noise.

4 Applications

Many of these parameter estimation techniques have been applied in retrospective studies of epidemiological events, such as SARS, influenza pandemics, as well as modern outbreaks such as Ebola and Zika. Though the problems they tackle are often very similar, there are many approaches taken, and compromises made.

Many studies tackle the nonconvexity and nonidentifiability of the trajectory matching methods head-on. One study that utilises trajectory matching, for example, is D'Silva and Eisenberg [29], where the spatial spread of Ebola in West Africa in 2014-2016 is modelled with a gravity model, as introduced in Section 2.2.2.

[INCOMPLETE]

There are also Bayesian approaches to these problems. For example, Stan, a Hamiltonia Monte Carlo (HMC) implementation, was used in the analysis of a 1978 influenza outbreak [45].
[INCOMPLETE]

Perhaps most pertinently, Hooker *et al.* [38] applied the generalised profiling technique to the Ontario measles cases over a period of fifty years from 1939 to 1989. In this paper, a standard SEIR model with time-dependent transmissivity parameter is fitted. They also introduce the concept of forward cross-validation as an option ofr determining the model regularisation parameter, λ. This method requires the computation of the error between the recovered (spline) trajectories and the trajectories predicted by the recovered parameters for each value of λ, at a subset of time values of the time domain t. Within the paper, they also very briefly touch on aspects of practical identifiability, noting the flat profile of the objective function over a parameter of a forcing term; other parameters are not discussed.

5 Future Directions

5.1 Characterising Contact Networks

As mentioned in Section 2.2, the modelling of contact networks is still not a mature field. Data on contact networks is difficult to acquire, as it requires relatively invasive methods to acquire fine-grain data. Recently,

5.2 Prior Knowledge for Estimation

How to synthesise knowledge on disease presentation with data properly (correct prior models).

How to use results from previous analyses

How to validate analyses, benchmarks

5.3 Model Inadequacy

Methods of characterising or even detecting model inadequacy

5.4 Estimation of Spatial Models

Computing approximations to the infinite dimensional parameter space that is a spatial solution is hard. It's also tough to characterise movement. Bit lost actually

5.5 Tools For Identifiability

Identifiability has been done for ODE systems. PRofile likelihood tools not widely available, or fast.

6 Summary

Epidemiological modelling is important for the control and prevention of

References

[1] R. Ross, *The prevention of malaria*, 2nd. London: Murray, 1911.

[2] W. O. Kermack and A. G. McKendrick, "A Contribution to the Mathematical Theory of Epidemics," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 115, no. 772, pp. 700–721, Aug. 1927, ISSN: 1364-5021. DOI: 10.1098/rspa.1927.0118.

[3] V. Capasso, *Mathematical Structures of Epidemic Systems*, ser. Lecture Notes in Biomathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1993, vol. 97, ISBN: 978-3-540-56526-0. DOI: 10.1007/978-3-540-70514-7.

[4] J. Heffernan and M. Keeling, "Implications of vaccination and waning immunity," *Proceedings of the Royal Society B: Biological Sciences*, vol. 276, no. 1664, pp. 2071–2080, Jun. 2009, ISSN: 0962-8452. DOI: 10.1098/rspb.2009.0057.

[5] J. Dushoff, J. B. Plotkin, S. A. Levin, and D. J. D. Earn, "Dynamical resonance can account for seasonality of influenza epidemics.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 48, pp. 16 915–6, Nov. 2004, ISSN: 0027-8424. DOI: 10.1073/pnas.0407293101.

[6] J. Shaman and M. Kohn, "Absolute humidity modulates influenza survival, transmission, and seasonality.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 9, pp. 3243–8, Mar. 2009, ISSN: 1091-6490. DOI: 10.1073/pnas.0806852106.

[7] D. Tompkins, A. Brock, G. Jones, G. Mcbride, A. Tait, J. Benschop, J. Marshall, N. French, S. Harper, A. Parshotam, W. Ye, D. Anderson, C. Macleod, and D. Slaney, "Modelling the Impacts of Climate Change On Infectious Diseases in New Zealand Health Analysis & Information For Action (HAIFA)," Environmental Science and Research Limited, Porirua, Tech. Rep., 2012, p. 64.

[8] F. Brauer and C. Castillo-Chavez, *Mathematical Models for Communicable Diseases*, 84. 2013, ISBN: 9781611972412. DOI: 10.1137/1.9781611972429.

[9] R. Yaari, I. Dattner, and A. Huppert, "A two-stage approach for estimating the parameters of an age-group epidemic model from incidence data," *Statistical Methods in Medical Research*, vol. 27, no. 7, pp. 1999–2014, Jul. 2018, ISSN: 0962-2802. DOI: 10.1177/0962280217746443.

[10] O. Diekmann, J. Heesterbeek, and J. Metz, "On the definition and the computation of the basic reproduction ratio R 0 in models for infectious diseases in heterogeneous populations," *Journal of Mathematical Biology*, vol. 28, no. 4, pp. 365–382, Jun. 1990, ISSN: 0303-6812. DOI: 10.1007/BF00178324.

[11] M. T. Meehan, D. Cocks, J. M. Trauer, and E. S. McBryde, *Coupled, multi-strain SIS, SIR and SIRS epidemic models*, 2016.

[12] M. T. Meehan, D. G. Cocks, J. M. Trauer, and E. S. McBryde, "Coupled, multi-strain epidemic models of mutating pathogens," *Mathematical Biosciences*, vol. 296, pp. 82–92, Feb. 2018, ISSN: 0025-5564. DOI: 10.1016/J.MBS.2017.12.006.

[13] Z. Feng, A. N. Hill, P. J. Smith, and J. W. Glasser, "An elaboration of theory about preventing outbreaks in homogeneous populations to include heterogeneity or preferential mixing," *Journal of Theoretical Biology*, vol. 386, pp. 177–187, Dec. 2015, ISSN: 0022-5193. DOI: 10.1016/J.JTBI.2015.09.006.

[14] L. J. S. Allen, "An introduction to stochastic epidemic models," in *Mathematical Epidemiology*, F. Brauer, P. van den Driessche, and J. Wu, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 81–130, ISBN: 978-3-540-78911-6. DOI: 10.1007/978-3-540-78911-6_3.

[15] L. J. S. Allen, S. R. Jang, and L.-I. Roeger, "Predicting population extinction or disease outbreaks with stochastic models," *Letters in Biomathematics*, vol. 4, no. 1, pp. 1–22, Jan. 2017, ISSN: 2373-7867. DOI: 10.1080/23737867.2016.1264870.

[16] J. Snow, *On the Mode of Communication of Cholera - John Snow - Google Books*, Second Edition, much Enlarged. London: John Churchill, New Burlington Street, 1855.

[17] J. Arino and P. Van Den Driessche, "A multi-city epidemic model," *Mathematical Population Studies*, vol. 10, no. 3, pp. 175–193, Jan. 2003, ISSN: 08898480. DOI: 10.1080/08898480306720.

[18] L. A. Rvachev and I. M. Longini, "A mathematical model for the global spread of influenza," *Mathematical Biosciences*, vol. 75, no. 1, pp. 3–22, Jul. 1985, ISSN: 0025-5564. DOI: 10.1016/0025-5564(85)90064-1.

[19] X. Li, H. Tian, D. Lai, and Z. Zhang, "Validation of the gravity model in predicting the global spread of influenza.," *International journal of environmental research and public health*, vol. 8, no. 8, pp. 3134–43, 2011, ISSN: 1660-4601. DOI: 10.3390/ijerph8083134.

[20] H. Heesterbeek, R. M. Anderson, V. Andreasen, S. Bansal, D. De Angelis, C. Dye, K. T. D. Eames, W. J. Edmunds, S. D. W. Frost, S. Funk, T. D. Hollingsworth, T. House, V. Isham, P. Klepac, J. Lessler, J. O. Lloyd-Smith, C. J. E. Metcalf, D. Mollison, L. Pellis, J. R. C. Pulliam, M. G. Roberts, C. Viboud, and I. N. I. I. Isaac Newton Institute IDD Collaboration, "Modeling infectious disease dynamics in the complex landscape of global health.," *Science (New York, N.Y.)*, vol. 347, no. 6227, aaa4339, Mar. 2015, ISSN: 1095-9203. DOI: 10.1126/science.aaa4339.

[21] D. G. Kendall, "Mathematical Models of The Spread of Infections, Mathematics and Computer Science in Biology and Medicine," in *Mathematics and Computer Science in Biology and Medicine*, London: H.M.S.O., 1965, pp. 218–225.

[22] J. D. Murray, *Mathematical biology. I, Introduction*, 2nd. Springer, 2002, p. 551, ISBN: 9780387224374.

[23] ——, *Mathematical biology. II Spatial models and biomedical applications*, 2nd. Springer, 2003, p. 811, ISBN: 9780387224381.

[24] L. Rass and J. Radcliffe, *Spatial deterministic epidemics*. American Mathematical Society, 2003, p. 261, ISBN: 9780821804995.

[25] J. M. Hyman and T. Laforce, "Modelling the Spread of Influenza among Cities," in *Bioterrorism : mathematical modelling applications to homeland security*, H. T. Banks and C. Castillo-Chavez, Eds., Philadelphia: Society for Industrial and Applied Mathematics, 2003, ch. 10, pp. 211–236, ISBN: 0-89871-549-0.

[26] J. Arino, J. R. Davis, D. Hartley, R. Jordan, J. M. Miller, and P. van den Driessche, "A multi-species epidemic model with spatial dynamics," *Mathematical Medicine and Biology: A Journal of the IMA*, vol. 22, no. 2, pp. 129–142, Jun. 2005, ISSN: 1477-8602. DOI: 10.1093/imammb/dqi003.

[27] D. Bichara, Y. Kang, C. Castillo-Chavez, R. Horan, and C. Perrings, "SIS and SIR Epidemic Models Under Virtual Dispersal," *Bulletin of Mathematical Biology*, vol. 77, no. 11, pp. 2004–2034, Nov. 2015, ISSN: 0092-8240. DOI: 10.1007/s11538-015-0113-5.

[28] C. Castillo-Chavez, D. Bichara, and B. R. Morin, "Perspectives on the role of mobility, behavior, and time scales in the spread of diseases," *Proceedings of the National Academy of Sciences*, vol. 113, no. 51, pp. 14582–14588, Dec. 2016, ISSN: 0027-8424. DOI: 10.1073/PNAS.1604994113.

[29] J. P. D'Silva and M. C. Eisenberg, "Modeling spatial invasion of Ebola in West Africa," *Journal of Theoretical Biology*, vol. 428, pp. 65–75, Sep. 2017. DOI: 10.1016/J.JTBI.2017.05.034.

[30] J. Poot, O. Alimi, M. P. Cameron, and D. C. Maré, "The gravity model of migration: the successful comeback of an ageing superstar in regional science," *Investigaciones Regionales*, no. 36, pp. 63–86, 2016.

[31] J. Truscott and N. M. Ferguson, "Evaluating the Adequacy of Gravity Models as a Description of Human Mobility for Epidemic Modelling," *PLoS Computational Biology*, vol. 8, no. 10, M. Pascual, Ed., e1002699, Oct. 2012, ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1002699.

[32] G. Chowell, C. Viboud, L. Simonsen, and S. M. Moghadas, "Characterizing the reproduction number of epidemics with early subexponential growth dynamics," *Journal of The Royal Society Interface*, vol. 13, no. 123, p. 20160659, Oct. 2016, ISSN: 1742-5689. DOI: 10.1098/rsif.2016.0659.

[33] G. Chowell, L. Sattenspiel, S. Bansal, and C. Viboud, "Mathematical models to characterize early epidemic growth: A review.," *Physics of life reviews*, vol. 18, pp. 66–97, 2016, ISSN: 1873-1457. DOI: 10.1016/j.plrev.2016.07.005.

[34] J. Ramsay and G. Hooker, *Dynamic Data Analysis*, ser. Springer Series in Statistics. New York, NY: Springer New York, 2017, ISBN: 978-1-4939-7188-6. DOI: 10.1007/978-1-4939-7190-9.

[35] A. Gábor and J. R. Banga, "Robust and efficient parameter estimation in dynamic models of biological systems.," *BMC systems biology*, vol. 9, p. 74, Oct. 2015, ISSN: 1752-0509. DOI: 10.1186/s12918-015-0219-2.

[36] N. Goeyvaerts, L. Willem, K. Van Kerckhove, Y. Vandendijck, G. Hanquet, P. Beutels, and N. Hens, "Estimating dynamic transmission model parameters for seasonal influenza by fitting to age and season-specific influenza-like illness incidence," *Epidemics*, vol. 13, pp. 1–9, Dec. 2015, ISSN: 1755-4365. DOI: 10.1016/J.EPIDEM.2015.04.002.

[37] I. Dattner, "A model-based initial guess for estimating parameters in systems of ordinary differential equations," *Biometrics*, vol. 71, no. 4, pp. 1176–1184, Dec. 2015, ISSN: 0006341X. DOI: 10.1111/biom.12348.

[38] G. Hooker, S. P. Ellner, L. D. V. Roditi, and D. J. D. Earn, "Parameterizing state-space models for infectious disease dynamics by generalized profiling: measles in Ontario.," *Journal of the Royal Society, Interface*, vol. 8, no. 60, pp. 961–74, Jul. 2011, ISSN: 1742-5662. DOI: 10.1098/rsif.2010.0412.

[39] G. Wahba, *Spline Models for Observational Data*. Society for Industrial and Applied Mathematics, Jan. 1990, ISBN: 978-0-89871-244-5. DOI: 10.1137/1.9781611970128.

[40] P. H. C. Eilers and B. D. Marx, "Flexible smoothing with B -splines and penalties," *Statistical Science*, vol. 11, no. 2, pp. 89–121, May 1996, ISSN: 0883-4237. DOI: 10.1214/ss/1038425655.

[41] J. O. Ramsay and B. W. Silverman, *Functional Data Analysis*, ser. Springer Series in Statistics. New York, NY: Springer New York, 2005, ISBN: 978-0-387-40080-8. DOI: 10.1007/b98888.

[42] J. O. Ramsay, G. Hooker, D. Campbell, and J. Cao, "Parameter estimation for differential equations: a generalized smoothing approach," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 69, no. 5, pp. 741–796, Nov. 2007, ISSN: 13697412. DOI: 10.1111/j.1467-9868.2007.00610.x.

[43] D. Campbell and O. Chkrebtii, "Maximum profile likelihood estimation of differential equation parameters through model based smoothing state estimates," *Mathematical Biosciences*, vol. 246, no. 2, pp. 283–292, Dec. 2013, ISSN: 0025-5564. DOI: 10.1016/J.MBS.2013.03.011.

[44] X. Xun, J. Cao, B. Mallick, A. Maity, and R. J. Carroll, "Parameter Estimation of Partial Differential Equation Models," *Journal of the American Statistical Association*, vol. 108, no. 503, pp. 1009–1020, Sep. 2013, ISSN: 0162-1459. DOI: 10.1080/01621459.2013.794730.

[45] A. Chatzilena, E. van Leeuwen, O. Ratmann, M. Baguelin, and N. Demiris, "Contemporary statistical inference for infectious disease models using Stan," Aug. 2019. arXiv: 1903.00423.

[46] I. Swameye, T. G. Muller, J. Timmer, O. Sandra, and U. Klingmuller, "Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 3, pp. 1028–33, Feb. 2003, ISSN: 0027-8424. DOI: 10.1073/pnas.0237333100.

[47] V. Raia, M. Schilling, M. Bohm, B. Hahn, A. Kowarsch, A. Raue, C. Sticht, S. Bohl, M. Saile, P. Moller, N. Gretz, J. Timmer, F. Theis, W.-D. Lehmann, P. Lichter, and U. Klingmuller, "Dynamic Mathematical Modeling of IL13-Induced Signaling in Hodgkin and Primary Mediastinal B-Cell Lymphoma Allows Prediction of Therapeutic Targets," *Cancer Research*, vol. 71, no. 3, pp. 693–704, Feb. 2011, ISSN: 0008-5472. DOI: 10.1158/0008-5472.CAN-10-2987.

[48] A. Smirnova, L. DeCamp, and H. Liu, "Inverse Problems and Ebola Virus Disease Using an Age of Infection Model," in *Mathematical and Statistical Modeling for Emerging and Re-emerging Infectious Diseases*, Cham: Springer International Publishing, 2016, pp. 103–121. DOI: 10.1007/978-3-319-40413-4_8.

[49] R. Muñoz-Tamayo, L. Puillet, J. B. Daniel, D. Sauvant, O. Martin, M. Taghipoor, and P. Blavy, "Review: To be or not to be an identifiable model. Is this a relevant question in animal science modelling?" *Animal*, vol. 12, no. 4, pp. 701–712, Apr. 2018, ISSN: 1751732X. DOI: 10.1017/S1751731117002774.

[50] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer, "Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood," *Bioinformatics*, vol. 25, no. 15, pp. 1923–1929, Aug. 2009, ISSN: 1460-2059. DOI: 10.1093/bioinformatics/btp358.

[51] N. Tuncer and T. T. Le, "Structural and practical identifiability analysis of outbreak models," *Mathematical Biosciences*, vol. 299, pp. 1–18, May 2018, ISSN: 0025-5564. DOI: 10.1016/J.MBS.2018.02.004.