Reformulating the susceptible-infectious-removed model in terms of the number of detected cases: well-posedness of the observational model

Article (Published Version)

Campillo-Funollet, Eduard, Wragg, Hayley, Van Yperen, James, Duong, Duc Lam and Madzvamuse, Anotida (2022) Reformulating the susceptible-infectious-removed model in terms of the number of detected cases: well-posedness of the observational model. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 380 (2233). a20210306 1-15. ISSN 1364-503X

This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/108185/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

http://sro.sussex.ac.uk
Reformulating the susceptible–infectious–removed model in terms of the number of detected cases: well-posedness of the observational model

Eduard Campillo-Funollet¹, Hayley Wragg²,³, James Van Yperen³, Duc-Lam Duong³,⁴ and Anotida Madzvamuse³,⁵,⁶

¹Department of Statistical Methodology and Applications, School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, Kent CT2 7PE, UK
²Department of Engineering Mathematics, School of Computer Science, Electrical and Electronic Engineering and Engineering Maths, University of Bristol, Bristol BS8 1TW, UK
³Department of Mathematics, School of Mathematical and Physical Sciences, University of Sussex, Brighton, East Sussex BN1 9QH, UK
⁴School of Engineering Science, LUT University, Lappeenranta 53850, Finland
⁵Department of Mathematics, University of Johannesburg, Johannesburg, South Africa
⁶University of British Columbia, Department of Mathematics, Vancouver, Canada

Compartmental models are popular in the mathematics of epidemiology for their simplicity and wide range of applications. Although they are typically solved as initial value problems for a system of ordinary differential equations, the observed data are typically akin to a boundary value-type problem: we observe some of the dependent variables at given
times, but we do not know the initial conditions. In this paper, we reformulate the classical susceptible–infectious–recovered system in terms of the number of detected positive infected cases at different times to yield what we term the observational model. We then prove the existence and uniqueness of a solution to the boundary value problem associated with the observational model and present a numerical algorithm to approximate the solution.

This article is part of the theme issue ‘Technical challenges of modelling real-life epidemics and examples of overcoming these’.

1. Introduction

At the beginning of any pandemic, mathematical modelling of infectious diseases is thrust into the limelight for governments across the globe to gain imperative forecasting pictures of the impending spread of the infectious disease, as demonstrated by the 2009–2010 swine flu pandemic [1–4], the 2015–2016 Zika virus epidemic [5–8] and the most recent COVID-19 pandemic [9–12]. Non-pharmaceutical interventions (NPIs) are interpreted by adjusting model parameters to allow for quick decision-making on the appropriate NPI measure(s) to be implemented. Understanding and predicting levels of infections in the community is the crux of mathematical modelling in epidemiology since it is these infections that drive the epidemic and also the mathematical models. However, getting accurate, reliable data to describe the current number of infections within the community is arguably the largest challenge for data-driven modelling. Not only is lack of data problematic when deriving and using mathematical models, the context and interpretation of the data and the parameter fitting process often comes with difficulties. In particular, there is often not a one-to-one correspondence between data and parameters or compartments in a model, especially when considering daily snapshots of an epidemic, for example.

In this work, we study the susceptible–infectious–removed (SIR) equations often accredited to the work of Kermack & McKendrick in 1927 [13]. In particular, we look to demonstrate how to use the number of detected cases directly in the model formulation and prove the existence and uniqueness of a solution to the resulting formulation. The SIR equations take the following form of a simple temporal dynamic system of ordinary differential equations, supported by non-negative initial conditions

\[
\begin{align*}
\dot{S} &= -\beta \frac{I}{N} S, \quad S(0) = S_0, \\
\dot{I} &= \beta \frac{I}{N} S - \gamma I, \quad I(0) = I_0, \\
\dot{R} &= \gamma I, \quad R(0) = R_0,
\end{align*}
\]

where the dot above the notation denotes the time derivative, see figure 1 for a schematic. As is standard for the SIR equations, we let \( N \) denote the total population being considered, \( \beta \) denote the average transmission rate and \( \gamma \) denote the average removal rate. The initial conditions are chosen so that \( S_0 + I_0 + R_0 = N \), which gives rise to \( S(t) + I(t) + R(t) = N \) for any \( t \geq 0 \). Moreover, it is standard to make the assumption that \( I_0 \geq 1 \), i.e. there is at least one infected individual at the onset of the infectious disease. We direct readers to the following literature to gain insight into the analysis and uses of the SIR equations and its family of compartmental models [14–17].

The ‘observational model’, as coined in [18], is a representation of a compartmental model, such as the SIR equations, described in terms of the parameters and compartments that are captured, or to be inferred, by the mathematical interpretation of the data. This observational model gives an intuitive understanding of what parameters can be identified and how the data affects different compartments in the model. Typically, data for infectious diseases come as a daily, or in general, periodic, snapshot of some of the states that individuals experience during
Figure 1. Schematic of the SIR compartmental model. We let $S$ denote the proportion of the total population $N$ who are susceptible to the infectious disease being studied. Susceptible individuals become infectious with the disease to form the $I$ subpopulation at a rate $\lambda(t)$, which represents the current average infection rate. The rate $\lambda(t)$ is the product of $\beta$, the average transmission rate, and the probability of meeting an infectious person $I(t)$ $N^{-1}$. Individuals in the $I(t)$ subpopulation then are removed at a rate $\gamma$ to form the removed subpopulation $R$. The description of the removed compartment depends on the nature of the infectious disease and the interpretation of the data available to be used. Typically, the term removed is interchanged with the term recovered. As is standard for epidemiological models of this nature, $\beta^{-1}$ denotes the average time between transmissions and $\gamma^{-1}$ denotes the average length of time before removal.

the infectious disease. In this study, we are looking to use the detected cases of an infectious disease, which we denote by $X_m$. In practical terms, we are assuming that our observation is a fraction $r \in (0, 1)$ of the total number of individuals leaving the infectious state $I$ over the time of the observation. This situation arises when the individual ceases to be infectious once they have been diagnosed, for example because the individual can avoid contact with susceptible individuals. We allow for the situation that we only observe a fraction $r$ of cases since there are situations where not all infectious individuals can be removed, such as false-negatives in testing regimes, bad individual compliance in behaviour or financial/resource restrictions on the amount of testing that can be done. In reality, $r$ is likely to be time-dependent since it can characterize aspects of public policy, however, we keep it constant in this study. In order to describe these observations in terms of the model parameters and compartments, we formally describe the removed compartment $R$ as those who are no longer mixing in the general population. Hence, for each observation of the number of detected cases $X_m, m = 0, \ldots, M$, we see that

$$X_m := r \gamma \int_{t_m}^{t_{m+1}} I(s) \, ds,$$  \hspace{1cm} (1.4)

where $[t_m, t_{m+1}]$ represents the time interval of the measurement of the data. This definition of the observed data for the number of detected cases of an infectious disease has been used in the past; see for instance [19].

Using (1.4) as the interpretation of the available data leads to the observational model

$$\dot{I} = \dot{I} \left( \frac{1}{I} - \beta \frac{1}{N} \right) - \beta \gamma \frac{I^2}{N},$$  \hspace{1cm} (1.5)

a second-order nonlinear differential equation equipped with non-local boundary conditions (1.4), see §3 for the detailed derivation. A significant part of this work is devoted to studying the boundary value problem (1.4) and (1.5). In contrast to initial value problems that typically arise when studying mathematical epidemiology, boundary value problems are much more difficult to deal with. For instance, it is well known that a small change in the boundary conditions may lead to significant changes in the behaviour of the solutions [20]. In fact, there is no general theory to the existence and uniqueness of a solution for boundary value problems. In most instances, we have to study the boundary value problem on a case-by-case basis, especially when the problem at hand is nonlinear, and/or the boundary conditions are atypical, which is the case with our problem as the boundary condition (1.4) is given in a non-local integral form. Non-local boundary conditions have been studied in the past in various forms, recently appearing in the study and application of fractional differential equations [21]. In our case, the non-local boundary conditions correspond to fixing the values of certain integrals of the solution over given time intervals, which is similar to how the average value is imposed on certain boundary value problems with Neumann boundary conditions to ensure uniqueness of the solution. Although this form is convenient for the analysis, we note that we can convert the integral conditions to
nonlinear, multi-point Robin-type conditions. This, however, is only possible due to the set-up of the observational model and thus the problem at hand, since the differential equation directly involves the data, and thus the boundary condition in some capacity. As far as we are aware, this is the first work to consider a boundary value problem with these specific integral-type boundary conditions. The key message of our study is on the transformation of an epidemiological initial value problem to a boundary value problem due to the appropriate form of the observed data, hence we refer readers to [20–26], and the references therein, for further information on methods and techniques that may be used to analyse certain classes of boundary value problems with different types of boundary conditions.

The main results of this publication can be summarized in the following two theorems:

**Theorem 1.1 (Existence).** Let \( \beta_c := r^{-1} \beta N^{-1} \). Given \( X_0, X_1 > 0 \) corresponding to (1.4), and given parameters \( \beta_c, \gamma > 0 \), there exists a solution \( I \in C^2(t_0, t_2) \) to (1.5) subject to (1.4).

**Theorem 1.2 (Uniqueness).** Let \( \beta_c := r^{-1} \beta N^{-1} \). Given \( X_0, X_1 > 0 \) corresponding to (1.4), and given parameters \( \beta_c, \gamma > 0 \), the solution to (1.5) subject to (1.4) is unique in \( C^2(t_0, t_2) \). Furthermore, the solution is smooth.

The structure of the paper is as follows: in §2, we motivate how the approach outlined in this work came about and why it is necessary, in §3, we derive the observational model and pose an equivalent boundary value problem to (1.5), in §4, we prove Theorem 1.1, in §5, we prove theorem 1.2 and in §6, we outline an algorithm for the numerical solution to the observational model with two examples.

## 2. Motivation

Since the onset of COVID-19, the disease caused by SARS-CoV-2, in December 2019 and the declaration of a pandemic by the World Health Organization in March 2020, governments and organizations across the globe have implemented unprecedented regional, national and international interventions to contain the spread of COVID-19 and the subsequent damage caused by it. Initially, NPIs were the only tool in the fight to contain the spread of the disease, forcing governments to introduce strict control measures on public mixing, the use of personal protective equipment, improvement of personal hygiene and the use of contact tracing to reduce the possibility of transmission throughout their countries, protecting those most vulnerable as well as maintaining control of the healthcare demand and capacity. The social-economical effects of these necessary decisions are still unknown, and it is speculated that these effects will still be felt for many more years to come [27–30]. The global picture is starting to look less bleak with the introduction of effective vaccines and vaccination policies allowing the relaxation of such strict NPIs, however with the spawn of new variants of COVID-19, such as the Alpha variant (lineage B.1.1.7) and more recently the Delta variant (lineage B.1.617.2), it is clear we are not out of the pandemic yet [31–34]. Indeed, now that ‘normality’ has returned, country borders have opened and tourism is on the increase, we in the UK are amidst an increase in daily positive cases, hospitalizations and deaths, while other European countries are tightening their COVID-19 restrictions or even going back into lockdowns.

In response to the pandemic in the Sussex region of the UK, local authority organizations in Sussex created a consortium that posed modelling questions of strategic operational significance to the Local Health Resilience Partnership covering East Sussex, West Sussex and Brighton & Hove. As a result, leaders from Public Health Intelligence in the local authority organizations approached the University of Sussex to undertake COVID-19 epidemiological modelling that is specific to Sussex as the national level modelling from the Scientific Advisory Group for Emergencies was not applicable at the regional, integrated care level. In response to their request, we developed a data-driven SIR-type mathematical model, which in practice has been used to answer public health questions based around healthcare demand and capacity, and mortuary capacity, such as the impact of second waves, future lockdowns and vaccination supply [18].
our early parameter estimation attempts, it became clear to us that the data provided were not compatible with the standard SIR-type formulation, in particular, the data did not provide an accurate estimation of all the initial conditions of the compartments of the model. Healthcare data typically come in one of three types: new daily data (such as admissions), current daily data (such as beds occupied) or cumulative daily data (such as total number of deaths). Due to the methods and constraints on data collection, data typically come as new daily data, i.e. a change in the state of an individual. For example, being admitted into hospital is a change of state from being infectious in the community to being in a hospital bed, whereby the infectious state and hospital bed state are compartments of a mathematical model, while the admitted component is the change between compartments. In this set-up, the admissions data are proportional to the infectious compartment, but not directly as it is an integral, so identifying the initial conditions of the compartments from the data becomes a real challenge, see [18] for more details.

Although the remit of this paper does not include parameter estimation results, we nevertheless describe these in order to demonstrate the issue of the estimation of initial conditions described above. Given that we have a system of nonlinear differential equations, the parameter estimation problem can be stated as a nonlinear least-squares (NLS) optimization problem [35]. Let \( d \) describe the observed data, \( \theta := [\beta, \gamma] \) describe the set of parameters we are estimating, \( x(\theta) := [S(\theta), I(\theta)] \) describe the set of solutions to (1.1) and (1.2) with \( f(x(\theta), \theta) \) describing the differential equations, and \( y(x, \theta) \) denote the function that maps the solutions \( x \) to the mathematical interpretation of the data. Let \( \mathcal{L}(d|\theta; y(x, \theta)) \) describe the likelihood function associated with the probability distribution assumption of the noise between \( d \) and \( y \). Then, the NLS optimization problem can be stated as

\[
\text{maximize}_{\theta} \quad \mathcal{L}(d|\theta; y(x, \theta))
\]

subject to \( \theta \in \mathcal{D}, \)
\[
\dot{x}(\theta) = f(x(\theta), \theta),
\]
\[
x(0) = x_0,
\]

where \( \mathcal{D} \) denotes the admissible parameter space. The NLS problem easily demonstrates the issue of the initial conditions faced: that even although parameters \( \theta \) can be estimated regardless of the initial conditions, they may not be representative of real life. This impacts the use of epidemiological models, that they may not have any accurate realistic forecasting capabilities. Moreover, it can easily be reasoned that \( \mathcal{L} \) is highly non-convex with respect to the parameters and in practice will therefore likely converge to solutions close to the initial guess, which may not be the global solution [36,37].

The standard approach in the literature to navigate the issue of estimating initial conditions is to restrict the parameters, for example, by fixing the basic reproductive number, and estimate when the epidemic began using the data provided, see for instance [38]. Fortunately for modellers in the UK, with regards to COVID-19, the disease was brought to us from outside our borders and so identifying ‘patient zero’ and tracing their connections becomes easier. However, with epidemiological data in particular, what is often not considered is that there is stochastic perturbation in the realization of the actual epidemic. The spread of an infectious disease is a stochastic process, and so the collection of data from an epidemic is just one realization of that process. Noise in the data presents itself mathematically as uncertainty in the parameters, while stochastic perturbation in the epidemic presents itself mathematically as uncertainty in both the initial conditions and the parameters. Variability due to stochastic perturbation is high at the onset of an epidemic, especially in a spatially complex and heterogeneous population not typically accounted for in modelling, which makes estimating the first infection and calibrating to data difficult. In [18], we took an alternative approach and estimated the initial conditions from the data at the onset of the first national lockdown, since this was the first available bit of data. In this instance, the initial conditions became part of the parameters we looked to estimate, however this proved challenging due to the nonlinearity of the NLS problem and the differences...
in parameter sensitivity between initial conditions and system parameters (such as $\beta$ and $\gamma$). This led us onto the problem we are trying to solve currently: Can we incorporate the data into the model formulation in a more natural way that overcomes the issue of accurate estimation of initial conditions? We believe that by rewriting the SIR equations as the observational model with boundary conditions, we can gain accurate estimates of parameters while bypassing issues with initial conditions. However, before approaching the parameter identifiability and estimation problems, we first need to demonstrate the non-trivial well-posedness of the observational model.

3. Derivation of the observational model

Before we derive the observational model, it is important we provide some standard analytical results of (1.1) and (1.2). It is a standard result in dynamical systems and phase space analysis that for positive initial conditions, we have $0 < S(t) < N$ and $0 < I(t) < N$ for $t \geq 0$. In particular, steady-state analysis tells us that

$$
\lim_{t \to \infty} (S(t), I(t), R(t)) = (S_\infty, 0, N - S_\infty),
$$

where $S_\infty > 0$. Due to (1.1), $S(t)$ is monotonically decreasing, which implies that

$$
\lim_{I \to 0} \frac{\dot{I}(t)}{I(t)} = \lim_{I \to 0} \frac{\beta S(t)}{N} - \gamma = \beta \frac{S_\infty}{N} - \gamma. \tag{3.1}
$$

In addition, since the right-hand sides of (1.1) and (1.2) are smooth functions, their solutions are infinitely differentiable (see, for instance, [39]), which allows us to make the necessary steps to derive the observational model. Turning our attention to the data (1.4), for the analysis, we assume that, for any $m = 0, \ldots, M$,

$$
0 < X_m < \infty.
$$

That is to say that we always observe a strictly positive, finite number of cases since we are not considering any observation error or model error in this study, i.e. our observations are exact. Indeed, in this situation, the case $X_m = 0$ implies that there no more infectious cases in the population, and in consequence, there will be no more cases in future observations.

As described in [18], since the data (1.4) are described by the infectious compartment $I$, we need to rewrite equations (1.1) and (1.2) to be purely in terms of $I$ in order to derive the observational model. To this end, we note three properties of equations (1.1) and (1.2), namely

$$
\dot{S} + \dot{I} = -\gamma I, \tag{3.2}
$$

which, by differentiating, immediately gives us

$$
\ddot{S} + \ddot{I} = -\gamma \dot{I}. \tag{3.3}
$$

Moreover, by differentiating (1.1), we see that

$$
\ddot{S} = \dot{S} \left( \frac{\dot{I}}{I} - \frac{\beta I}{N} \right) \tag{3.4}
$$

and hence, inserting (3.2) and (3.3) into (3.4), the observational model takes the form

$$
\ddot{I} = \dot{I} \left( \frac{\dot{I}}{I} - \frac{\beta I}{N} \right) - \beta \gamma \frac{I^2}{N},
$$

as seen in (1.5). One notes that (3.4) and thus (1.5) are well defined for $t \geq 0$ due to (3.1). For the purposes of this paper, we continue to transform (1.5) to facilitate the analysis. By denoting
\( \beta_e := r^{-1} \beta N^{-1} \) and setting \( z(t) = \ln(\gamma \ I(t)) , \) (1.5) is equivalent to
\[
\ddot{z} = -\beta_e e^z \left( \frac{\dot{z}}{\gamma} + 1 \right) ,
\] (3.5)
while the data (1.4) are equivalent to
\[
X_m = \int_{t_n}^{t_{n+1}} e^{z(s)} \, ds .
\] (3.6)
In the subsequent sections, we will prove the following two theorems, which are respectively equivalent to Theorems 1.1 and 1.2 due to the continuous bijective transformation between \( z \) and \( I . \)

**Theorem 3.1 (Existence).** Given \( X_0, X_1 > 0 \), corresponding to (3.6), and given parameters \( \beta_e, \gamma > 0 \), there exists a solution \( z \in C^2(t_0, t_2) \) to (3.5) subject to (3.6).

**Theorem 3.2 (Uniqueness).** Given \( X_0, X_1 > 0 \) corresponding to (3.6), and given parameters \( \beta_e, \gamma > 0 \), the solution to (3.5) subject to (3.6) is unique in \( C^2(t_0, t_2) \). Furthermore, the solution is smooth.

**Remark.** One can observe that as \( t \) tends to 0, \( z \) becomes negatively unbounded. This is not impactful on the analysis due to the boundary condition (3.6), equivalently (1.4), always being positive. As we noted above, in the situation that (3.6) is zero, and upon assuming no noise in the data, this implies that the epidemic is over and thus this model is not relevant anymore.

**Remark.** One can go further to simplify (3.5) by taking \( w(\xi) = z(t) + \ln(\beta_e \gamma^{-2}) \), where \( \xi = \gamma t \), to get
\[
w'' = -e^w \left( w' + 1 \right) ,
\]
where the prime denotes the derivative with respect to \( \xi \), which leads to (3.6) satisfying
\[
X_m = \frac{\gamma}{\beta_e} \int_{t_n}^{t_{n+1}} e^{w(s)} \, ds .
\] (3.6)

**Remark.** To continue to demonstrate the issues with initial conditions, we can express \( S(0) \) and \( I(0) \) in terms of \( z \) and the parameters. By definition of \( z \), we obtain
\[
S(0) = N \frac{\dot{z}(0) + \gamma}{\beta} \] (3.7)
and
\[
I(0) = \frac{e^{z(0)}}{r \gamma} .
\] (3.8)
From the formulation of the observational model (3.5), we can see that a solution \( z \) is defined only in terms of \( \beta_e \) and \( \gamma \). Assuming that \( N \) is known, due to the definition of \( \beta_e \), the pair \( (\beta, r) \) cannot be uniquely identified from the data; in other words, for a given \( \beta_e \), any pair \( (\beta, r) \) satisfying \( \beta_e = r^{-1} \beta N^{-1} \) will produce exactly the same solution. This implies that in order to simulate future predictions with the original SIR formulation, one would need estimates for \( r \), as well as the initial conditions. Indeed, from (3.7), we conclude that it is not possible to identify \( S(0) \) unless both \( \beta \) and \( r \) are known; in particular, since it is not possible to identify both \( \beta \) and \( r \) from the data, it is not possible to recover the initial conditions without an independent estimate of the under-reporting coefficient \( r \). By contrast, an advantage of the formulation of the observational model is that it will allow us to compute the solution \( z \), and predict future observations \( X_m \), as long as we can fit the model to identify the value of \( \beta_e \). Note also that, without an independent estimate on \( r \), we can only infer the ratio between the initial number of infectious individuals to the initial number of susceptible individuals since
\[
\frac{I(0)}{S(0)} = \beta_e \frac{e^{z(0)}}{\gamma \dot{z}(0) + \gamma} .
\]
4. Existence

The proof of theorem 3.1 is based on the Leray–Schauder fixed point theorem [40]. We shall show that the solutions to (3.5) and (3.6) satisfy a priori bounds, and we will use them to show compactness of a suitable operator.

**Lemma 4.1.** Let $X_0 > 0$ and $X_1 > 0$. Suppose there exists a solution $z$ to (3.5) and (3.6) with data $X_0$ and $X_1$. Then, there exists some constant $C$ such that $|z(t)| \leq C$, for $t \in [t_0, t_2]$. Moreover, $C$ is only in terms of the parameters and the data.

**Proof.** Let $\bar{\beta}_e := \beta_e \gamma^{-1}$ and let $\tau \in [t_0, t_2]$. Integrating (3.5) over $(t_0, \tau)$, we have

$$
\dot{z}(\tau) - \dot{z}(t_0) = \bar{\beta}_e (e^{z(t_0)} - e^{\gamma \tau}) - \beta_\gamma \int_{t_0}^{\tau} e^{\gamma \tau} \, ds.
$$

Let $t \in [t_0, t_2]$ and let $\delta_0(t) := t - t_0$. Then, by integrating with respect to $\tau \in (t_0, t)$, we have

$$
z(t) - z(t_0) - \delta_0(t) \dot{z}(t) = \bar{\beta}_e \left( \delta_0(t) e^{z(t_0)} - \int_{t_0}^{t} e^{\gamma \tau} \, d\tau \right) - \beta_\gamma \int_{t_0}^{t} e^{\gamma \tau} \, ds \, d\tau.
$$

Since the exponential function is positive everywhere, we see that

$$
0 \leq \int_{t_0}^{t} e^{\gamma \tau} \, d\tau \leq \int_{t_0}^{t_2} e^{\gamma \tau} \, d\tau = X_0 + X_1,
$$

and consequently

$$
0 \leq \int_{t_0}^{t} \int_{t_0}^{\tau} e^{\gamma \tau} \, ds \, d\tau \leq \delta_0(t)(X_0 + X_1).
$$

Hence, we see that there exist bounds

$$
z_{0,l}(t) \leq z(t) \leq z_{0,u}(t),
$$

for $t \in [t_0, t_2]$, where

$$
z_{0,l}(t) := z(t_0) + \delta_0(t)(\dot{z}(t_0) + \bar{\beta}_e e^{z(t_0)}) - \beta_\gamma (1 + \gamma \delta_0(t))(X_0 + X_1)
$$

and

$$
z_{0,u}(t) := z(t_0) + \delta_0(t)(\dot{z}(t_0) + \bar{\beta}_e e^{z(t_0)}).
$$

These bounds depend on $z$ and $\dot{z}$ through their values at $t_0$. Since we are interested in uniform bounds for any solution $z$, we need to show that $z(t_0)$ and $\dot{z}(t_0)$ are bounded in terms of the data $X_0$ and $X_1$. By exponentiation and integration (4.1) over $(t_0, t_2)$, we have

$$
\int_{t_0}^{t_2} e^{z_{0,l}(t)} \, dt \leq (X_0 + X_1) \leq \int_{t_0}^{t_2} e^{z_{0,u}(t)} \, dt,
$$

and hence, by taking (4.2) into account,

$$
\int_{t_0}^{t_2} e^{z_{0,l}(t)} \, dt = e^{z(t_0)} - \bar{\beta}_e (X_0 + X_1) \int_{t_0}^{t_2} e^{\delta_0(t)(\dot{z}(t_0) + \bar{\beta}_e e^{z(t_0)} - \beta_\gamma (X_0 + X_1))} \, dt
$$

$$
= \frac{e^{z(t_0) - \bar{\beta}_e (X_0 + X_1)}}{\dot{z}(t_0) + \bar{\beta}_e e^{z(t_0)} - \beta_\gamma (X_0 + X_1)} \left( e^{\delta_0(t)(\dot{z}(t_0) + \bar{\beta}_e e^{z(t_0)} - \beta_\gamma (X_0 + X_1))} - 1 \right).
$$

To show that $z(t_0)$ and $\dot{z}(t_0)$ are bounded, we proceed by contradiction. Let $\alpha_0 := e^{z(t_0)}$ and $\sigma_0 := \dot{z}(t_0)$, and let

$$
f_{0,l}(\alpha_0, \sigma_0) := \frac{\alpha_0 e^{-\bar{\beta}_e (X_0 + X_1)}}{\alpha_0 + \bar{\beta}_e \sigma_0 - \beta_\gamma (X_0 + X_1)} \left( e^{\delta_0(t)(\sigma_0 + \bar{\beta}_e e^{\sigma_0} - \beta_\gamma (X_0 + X_1))} - 1 \right).
$$

A standard application of L’Hôpital’s rule yields

$$
\lim_{\alpha_0 \to \infty} f_{0,l}(\alpha_0, \sigma_0) = \infty \quad \text{and} \quad \lim_{\sigma_0 \to \infty} f_{0,l}(\alpha_0, \sigma_0) = \infty.
$$
where we have fixed $\sigma_0$ in the first limit and fixed $a_0$ in the second limit, and similarly

\[
\lim_{a_0, \sigma_0 \to \infty} f_{0,j} = \infty.
\]

Since (4.4) implies boundedness of $f_{0,j}$, we conclude $a_0$ and $\sigma_0$ cannot be arbitrarily large and in consequence, $z(t_0)$ and $\dot{z}(t_0)$ are bounded from above. Using a similar argument with the right-hand side bound of (4.4) and taking the limits $a_0 \to 0$ and $\sigma_0 \to -\infty$, we can conclude that $z(t_0)$ and $\dot{z}(t_0)$ are also bounded from below.

Therefore, the bounds (4.2) and (4.3) are independent of the solution $z$, and depend only on the parameters and the data.

We can now proceed to show that $\dot{z}$ is also uniformly bounded in $[t_0, t_2]$.

**Lemma 4.2.** Let $X_0 > 0$ and $X_1 > 0$. Suppose there exists a solution $z$ to (3.5) and (3.6) with data $X_0$ and $X_1$. Then, there exists some constant $C'$ such that $|\dot{z}(t)| \leq C'$, for $t \in [t_0, t_2]$. Moreover, $C'$ is only in terms of the parameters and the data.

**Proof.** Using the notation introduced in the proof of lemma 4.1, we integrate (3.5) over $(t_0, t)$, where $t \in [t_0, t_2]$, to obtain

\[
\dot{z}(t) = \dot{z}(t_0) - \frac{\beta e}{e^{\alpha - 1}} (e^{\alpha z(t)} - e^{\alpha z(t_0)}) - \beta e \int_{t_0}^{t} e^{\alpha z(s)} ds.
\]

We can now use the uniform bounds on $z$, from lemma 4.1, together with the bounds on $z(t_0)$ and $\dot{z}(t_0)$ that we obtained in the proof of lemma 4.1, to bound the right-hand side of (4.5).

We can proceed using equations (3.5) and (3.6) and the bounds on $z$ and $\dot{z}$ to obtain uniform bounds on the second derivative, which in turns allows us to obtain uniform bounds for the third derivative by using them in (3.5). Indeed, we can bootstrap this argument to obtain uniform bounds up to the $n$th derivative, for finite $n$. We summarize this in the following corollary.

**Corollary 4.3.** Let $n = 0, 1, \ldots, N$, with $N < \infty$, and let $X_0 > 0$ and $X_1 > 0$. Suppose there exists a solution $z$ to (3.5) and (3.6) with data $X_0$ and $X_1$. Then, there exists some constant $C_n$, depending only on $n$, the data and the parameters, such that $|z^{(n)}(t)| < C_n$ for all $t \in [t_0, t_2]$.

We can now use the uniform bounds on the solutions to study a suitable operator that will allow us to define the solutions of (3.5) and (3.6) as a fixed point. Let

\[
\mathcal{F}[z](t) := a_z + b_z t - \int_{t_0}^{t} \int_{t_0}^{s} \beta e e^{\alpha z(s)} \left( \frac{\dot{z}(s)}{\gamma} + 1 \right) ds d\tau,
\]

where $a_z, b_z$ depend on $z$ and are uniquely determined under the constraints

\[
\int_{t_0}^{t_1} e^{\mathcal{F}[z](t)} dt = X_0 \quad \text{and} \quad \int_{t_1}^{t_2} e^{\mathcal{F}[z](t)} dt = X_1.
\]

We note importantly that a fixed point of $\mathcal{F}$ is a solution to (3.5)–(3.6). Therefore, our goal is to apply the Leray–Schauder fixed point theorem to $\mathcal{F}$ to show the existence of a solution in $C^2(t_0, t_2)$. We will first show that $\mathcal{F}$ is a compact operator.

**Lemma 4.4.** Let $\mathcal{F}$ be defined as in (4.6). Then $\mathcal{F}$ is a compact operator in $C^2(t_0, t_2)$.

**Proof.** Let $(z_n)_{n \in \mathbb{N}_0}$ be a sequence of functions bounded in $C^2(t_0, t_2)$. We see that since

\[
(F[z_n])^{(3)} = -\beta e z_n e^{z_n} \left( \frac{z_n}{\gamma} + 1 \right) - \beta e e^{z_n},
\]

we have $(F[z_n])^{(3)}$ is also uniformly bounded, since $(z_n)_{n}$ is uniformly bounded in $C^2(t_0, t_2)$. Therefore, the Arzela–Ascoli theorem implies that $\mathcal{F}$ is a compact operator in $C^2(t_0, t_2)$. We note that to apply the Arzela–Ascoli theorem in $C^2(t_0, t_2)$ suffices to show uniform boundedness of the third derivative, since then we also have equicontinuity of the second derivative; this is analogous to using uniform boundedness of the first derivative to show equicontinuity of the sequence of functions.
With this set-up, in particular that $\mathcal{F}$ is compact in $C^2(t_0, t_2)$, we can now finish the proof of theorem 3.1.

**Proof of theorem 3.1.** In order to apply the Leray–Schauder fixed point theory, we now need to show that the set

$$
\{ z \in C^2(t_0, t_2) : z = \kappa \mathcal{F}[z] \}
$$

is bounded for all $\kappa \in [0, 1]$. It is easy to see, by linearity of the derivative and the integral, that a fixed point of $\kappa \mathcal{F}$ corresponds to a fixed point of $\mathcal{F}$ with a change of parameters, namely, by replacing $\beta e$ by $\kappa \beta e$. Therefore, any function $z$ in (4.8) is a solution to (3.5) and (3.6) with a suitable parameter, and as shown in lemma 4.1, is bounded in $C^2(t_0, t_2)$. Thus the Leray–Schauder fixed point theorem implies the existence of a fixed point for $\mathcal{F}$, which, as already mentioned, corresponds to a solution of (3.5)–(3.6).

5. Uniqueness

Now that we have proven that there exists a solution to (3.5)–(3.6), we now want to demonstrate the uniqueness of a solution to (3.5)–(3.6). We employ a standard contradiction argument for boundary value problems by showing that two different solutions, which satisfy the same boundary conditions, must be the same.

**Proof of theorem 3.2.** Let $x(t)$ and $y(t)$ be two solutions to (3.5) and (3.6) satisfying the same data conditions, that is, for $m = 0, 1$ we have

$$
X_m = \int_{t_m}^{t_{m+1}} e^{\alpha(s)} \, ds = \int_{t_m}^{t_{m+1}} e^{\beta(s)} \, ds.
$$

Consequently, this means that

$$
\int_{t_m}^{t_{m+1}} e^{\alpha(s)} - e^{\beta(s)} \, ds = 0. \quad (5.1)
$$

Let $D(t) := x(t) - y(t)$ and $E(t) := e^{\alpha(t)} - e^{\beta(t)}$. By taking the difference of (3.5) for each solution, we have

$$
\ddot{D} = -\beta e \left( E + \frac{\dot{E}}{\gamma} \right). \quad (5.2)
$$

To simplify the presentation, we set $\bar{\beta}_e := \beta_e \gamma^{-1}$ and denote

$$
F(t) := \dot{D}(t) + \bar{\beta}_e E(t) - \dot{D}(t_0) - \bar{\beta}_e E(t_0).
$$

Using this notation, (5.2) is equivalent to

$$
\ddot{F} + \beta_e E = 0. \quad (5.3)
$$

Let $t \in [t_0, t_2]$. Integrating (5.3) with respect to $s \in (t_0, t)$ gives

$$
F(t) + \beta_e \int_{t_0}^{t} E(s) \, ds = 0. \quad (5.4)
$$

Moreover, by (5.1) and (5.4), one can see that

$$
F(t_1) = F(t_2) = 0. \quad (5.5)
$$

We now proceed by contradiction to show that $D(t_0) = 0$. Without loss of generality, assume that $D(t_0) > 0$. This gives us that $E(t_0) > 0$ by definition and $\dot{F}(t_0) < 0$ by (5.3). Considering this, (5.5) and the fact that $x, y \in C^2(t_0, t_2)$, we have that there exists $\tau \in (t_0, t_1]$ such that $F(\tau) = 0$ and $F(t) < 0$...
for all \( t \in (t_0, \tau) \). In particular, this means that, by integrating \( F \) with respect to \( s \in (t_0, \tau) \), we have
\[
\int_{t_0}^{\tau} F(s) \, ds < 0
\]
(5.6)
and, setting \( t = \tau \) in (5.4), we have
\[
\int_{t_0}^{\tau} E(s) \, ds = 0.
\]
(5.7)

Then, by integrating (5.4) with respect to \( t \in (t_0, \tau) \) and using (5.6), we obtain
\[
0 = \int_{t_0}^{\tau} F(t) \, dt + \beta_c \int_{t_0}^{\tau} \int_{t_0}^{t} E(s) \, ds \, dt < \beta_c \int_{t_0}^{\tau} \int_{t_0}^{t} E(s) \, ds \, dt = 0,
\]
where the last equality is a consequence of Fubini’s theorem and (5.7). Therefore, we have a contradiction. One notices that, using the same approach, we can also conclude that \( D(t_1) = 0 \) due to (5.5). By integrating (5.4) with respect to \( t \in (t_0, t_1) \), we have that
\[
D(t_1) + \beta_c \int_{t_0}^{t_1} E(t) \, dt + \beta_c \int_{t_0}^{t_1} \int_{t_0}^{t} E(s) \, ds \, dt = D(t_0) + (t_1 - t_0)(\bar{D}(t_0) + \bar{D}(t_0)),
\]
which, by noting that \( D(t_0) = 0 \) implies that \( E(t_0) = 0 \) and using (5.1), gives us
\[
D(t_0) = 0.
\]

Standard ODE theory implies that, since \( x(t_0) = y(t_0) \), \( \dot{x}(t_0) = \dot{y}(t_0) \) and \( x, y \in C^2(t_0, t_2) \), we have
\[
x(t) = y(t)
\]
for all \( t \in [t_0, t_2] \). \( \blacksquare \)

6. Numerical examples

Now that we have demonstrated existence and uniqueness, we can build a numerical algorithm to approximate the solution to (3.5)–(3.6). We set up the numerical problem as follows: given \( X_0 \), \( X_1 \), \( \beta_c \) and \( \gamma \), we look to find the function \( z \) that satisfies (3.5)–(3.6). We use the shooting method to numerically approximate the solution to (3.5), see for example [41] for details. We use a superscript 0 to denote the initial approximation to \( z \) and a superscript \( \varepsilon \) to denote the converged solution from the shooting method.

For our first example, we look at the case where the number of detected cases is reducing. We set \( \beta_c = 8 \times 10^{-4} \), \( \gamma = 1 \) and use the data \( X_0 = 100 \) and \( X_1 = 50 \) with \( t_0 = 0 \), \( t_1 = 1 \) and \( t_2 = 2 \). For the shooting method, we use an initial guess of \( z^0(0) = 6 \), \( \dot{z}^0(0) = -1 \), depicted by the dashed blue line in figure 2a, which results in
\[
\int_{0}^{1} e^{z^0(s)} \, ds \approx 255.00079468571883 \quad \text{and} \quad \int_{1}^{2} e^{z^0(s)} \, ds \approx 93.80954898172891.
\]
The resulting initial conditions from the shooting method are \( z^\varepsilon(0) = 4.92579565 \) and \( \dot{z}^\varepsilon(0) = -0.66986594 \), with the resulting function \( z^\varepsilon \) depicted by the continuous orange line in figure 2a. Even although we can typically gain knowledge of the total population \( N \), through for example a population census, as previously mentioned we cannot extract \( \beta \) from \( \beta_c \) due to the under-reporting parameter \( r \). However, for the purpose of demonstration, we fix \( r = 0.75 \) and \( N = 1000 \) to allow us to demonstrate \( I^0 \) and \( I^\varepsilon \), the number of infectious people transformed from \( z^0 \) and \( z^\varepsilon \), respectively. Figure 2b depicts the transformed initial guess \( I^0 \) as the dashed blue line and depicts the transformed converged solution \( I^\varepsilon \) as the continuous orange line.

For our second example, we look at the case where the number of detected cases is increasing. We set \( \beta_c = 2 \times 10^{-3} \), \( \gamma = 1 \) and use the data \( X_0 = 50 \) and \( X_1 = 200 \) with \( t_0 = 0 \), \( t_1 = 1 \) and \( t_2 = 2 \). For the shooting method, we use an initial guess of \( z^0(0) = 6 \) and \( \dot{z}^0(0) = 1 \), depicted by the dashed...
Figure 2. Demonstration of the numerical approach to approximate the solutions to (3.5) and (3.6) in the case of the number of detected cases is reducing. (a) Comparison of the initial guess for $z$ and the converged numerical approximation for $z^\varepsilon$. (b) Comparison of the transformed initial guess for $I$ and the transformed converged numerical approximation for $I^\varepsilon$. (Online version in colour.)

Figure 3. Demonstration of the numerical approach to approximate the solutions to (3.5) and (3.6) in the case of the number of detected cases is increasing. (a) Comparison of the initial guess for $z$ and the converged numerical approximation for $z^\varepsilon$. (b) Comparison of the transformed initial guess for $I$ and the transformed converged numerical approximation for $I^\varepsilon$. (Online version in colour.)

The resulting initial conditions from the shooting method are $z^\varepsilon(0) = 2.9481878$ and $\dot{z}^\varepsilon(0) = 1.75958188$, with the resulting function $z^\varepsilon$ depicted as the continuous orange line in figure 3a. Again, we fix $r = 0.75$ and $N = 1000$ to allow us to demonstrate $I^0$ and $I^\varepsilon$ in figure 3b, whereby the transformed initial guess $I^0$ is depicted by the dashed blue line and the transformed converged solution $I^\varepsilon$ is depicted by the continuous orange line.

7. Conclusion

In this publication, we have demonstrated how to derive the observational model given data that describes the infectious compartment $I$, shown that a solution to this formulation exists and that
the solution is unique. We have then gone on to demonstrate an algorithm to approximate the solution to the nonlinear observational model. While it is standard these days for more complex and descriptive compartmental models to be analysed and used in practice, this study provides a first illustration of the analysis for the novel boundary value approach, which we intend to build upon to combat the issues outlined in §§2 and 3. It is not hard to reason that using two data points is not enough to provide accurate forecasting capabilities for the short and medium term, but this result combined with parameter estimation with more data points and a model selection approach will shed light on how one can use this approach.

Now that we have guaranteed that the solution to the observational model is unique, given some parameters, we next look to answer the questions of the identifiability of the parameters from the data and how to appropriately deal with noise in the data. That is to say, how much data do we need to be able to uniquely identify the parameters \( \beta_e \) and \( \gamma \) as well as \( z \)? These are questions of particular importance as understanding what we can identify, and the accuracy of the estimations, helps us to understand the scenario-based forecasting capabilities of the mathematical model. Understanding this and therefore the underlying assumptions needed to produce forecasting results, given the limelight of compartmental models in recent history, will allow for a confident and trustworthy exchange of knowledge between mathematical modellers and those with research questions on infectious diseases, such as local government or healthcare providers.

**Data accessibility.** This article has no additional data.

**Authors’ contributions.** E.C.-F.: conceptualization, formal analysis, funding acquisition, investigation, methodology, validation, visualization, writing—original draft, writing—review and editing; H.W.: formal analysis, investigation, validation, visualization, writing—review and editing; J.V.Y.: conceptualization, formal analysis, funding acquisition, investigation, methodology, supervision, validation, visualization, writing—original draft, writing - review and editing; D.-L.D.: formal analysis, investigation, methodology, validation, visualization, writing—review and editing; A.M.: conceptualization, funding acquisition, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

**Conflict of interest declaration.** We declare that we have no competing interests.

**Funding.** H.W. was supported by the Higher Education Innovation Fund through the University of Sussex, J.V.Y. and D.-L.D. were supported by Brighton and Hove City Council, East and West Sussex County Council and the NHS Sussex Commissioners, and E.C.-F. was supported by the Wellcome Trust grant no. 204833/Z/16/Z. This work was partly supported by the Global Challenges Research Fund through the Engineering and Physical Sciences Research Council grant no. EP/T00410X/1: UK-Africa Postgraduate Advanced Study Institute in Mathematical Sciences (A.M. and E.C.-F). A.M.’s work was partially funded by grants from the Health Foundation (1902431) and the NIHR (NIHR133761) and by an individual grant from Dr Perry James (Jim) Browne Research Centre on Mathematics and its Applications (University of Sussex). A.M. is a Royal Society Wolfson Research Merit Award Holder funded generously by the Wolfson Foundation.

**Acknowledgements.** All authors acknowledge the continued support and collaboration of Brighton and Hove City Council, East and West Sussex County Council and the NHS Sussex Commissioners and thank them for the opportunity for past collaborations that led to the conceiving of this manuscript.

**References**

1. Cauchemez S, Ferguson NM, Wachtel C, Tegnell A, Saour G, Duncan B, Nicoll A. 2009 Closure of schools during an influenza pandemic. *Lancet Infect. Dis.* 9, 473–481. (doi:10.1016/S1473-3099(09)70176-8)
2. Fraser C. 2009 *et al.* Pandemic potential of a strain of influenza A (H1N1): early findings. *Science* 325, 1557–1561. (doi:10.1126/science.1176062)
3. Fineberg HV. 2014 Pandemic preparedness and response—lessons from the H1N1 influenza of 2009. *N. Engl. J. Med.* 370, 1335–1342. (doi:10.1056/NEJMr1208802)
4. Yang Y, Sugimoto JD, Halloran ME, Basta NE, Chao DL, Matrajt L, Potter G, Kenah E, Longini IM. 2009 The transmissibility and control of pandemic influenza A (H1N1) virus. *Science* 326, 729–733. (doi:10.1126/science.1177373)
5. Asher J et al. 2017 Preliminary results of models to predict areas in the Americas with increased likelihood of Zika virus transmission in 2017. BioRxiv 187591 [Preprint]. (doi:10.1101/187591)
6. Ferguson NM, Cucunubá ZM, Dorigatti I, Nedjati–Gilani GL, Donnelly CA, Basáñez MG, Nouvellet P, Lessler J. 2016 Countering the Zika epidemic in Latin America. Science 353, 353–354. (doi:10.1126/science.aag0219)
7. Lowe R, Barcellos C, Brasil Pícia, Cruz O, Honório N, Kuper H, Carvalho M. 2018 The Zika virus epidemic in Brazil: from discovery to future implications. Int. J. Environ. Res. Public Health 15, 96. (doi:10.3390/ijerph15010096)
8. Zhang Q et al. 2017 Spread of Zika virus in the Americas. Proc. Natl Acad. Sci. USA 114, E4334–E4343. (doi:10.1073/pnas.1620161114)
9. Ferguson NM et al. 2020 Report 9: impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. See https://spiral.imperial.ac.uk:8443/handle/10044/1/77482. (doi:10.25561/77482)
10. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. 2020 Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science 368, 860–868. (doi:10.1126/science.abb5793)
11. Adam D. 2020 Special report: the simulations driving the world’s response to COVID-19. Nature 580, 316–318. (doi: 0.1038/d41586-020-01003-6)
12. Wu JT, Leung K, Leung GM. 2020 Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. Lancet 395, 689–697. (doi:10.1016/S0140-6736(20)30260-9)
13. Kermack WO, McKendrick AG. 1927 A contribution to the mathematical theory of epidemics. Proc. R. Soc. Lond. A 115, 700–721. (doi:10.1098/rspa.1927.0118)
14. Bailey NT. 1975 The mathematical theory of infectious diseases and its applications. High Wycombe, UK: Charles Griffin & Company LTD.
15. Brauer F, Castillo–Chavez C, Castillo–Chavez C. 2012 Mathematical models in population biology and epidemiology. New York, NY: Springer.
16. Diekmann O, Heesterbeek H, Britton T. 2012 Mathematical tools for understanding infectious disease dynamics. Princeton, NJ: Princeton University Press.
17. Hethcote HW. 2000 The mathematics of infectious diseases. SIAM Rev. 42, 599–653. (doi:10.1137/S0036144500371907)
18. Campillo-Funollet E et al. 2021 Predicting and forecasting the impact of local outbreaks of COVID-19: use of SEIR-D quantitative epidemiological modelling for healthcare demand and capacity. Int. J. Epidem. 50, 1103–1113. (doi:10.1093/ije/dyab106)
19. Magal P, Webb G. 2018 The parameter identification problem for SIR epidemic models: identifying unreported cases. J. Math. Biol. 77, 1629–1648. (doi:10.1007/s00285-017-1203-9)
20. Bernfeld SR, Lakshmikantham V. 1974 An introduction to nonlinear boundary value problems. New York, NY: Academic Press.
30. Nicola M, Alsaﬁ Z, Sohrabi C, Kerwan A, Al-Jabir A, Iosifidis C, Agha M, Agha R. 2020 The socio-economic implications of the coronavirus pandemic (COVID-19): a review. *Int. J. Surg.* 78, 185–193. (doi:10.1016/j.ijsu.2020.04.018)

31. Hodgson D, Flasche S, Jit M, Kucharski AJ. 2021 The potential for vaccinated-induced herd immunity against SARS-CoV-2 B.1.1.7 variant. *Euro Surveill.* 26, 2100428. (doi:10.2807/1560-7917.ES.2021.26.20.2100428)

32. Harris RJ, Hall JA, Zaidi A, Andrews NJ, Dunbar JK, Dahreba G. 2021 Impact of vaccination on household transmission of SARS-CoV-2 in England. See https://static.poder360.com.br/2021/04/public-health-1stdose-abr2021.pdf.

33. Davies NG et al. 2021 Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* 372, eabg3055. (doi:10.1126/science.abg3055)

34. Moore S, Hill EM, Tildesley MJ, Dyson L, Keeling MJ. 2021 Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study. *Lancet Infect. Dis.* 21, 793–802. (doi:10.1016/S1473-3099(21)00143-2)

35. Stortelder WJ. 1996 Parameter estimation in dynamic systems. *Math. Comput. Simul.* 42, 135–142. (doi:10.1016/0378-4754(95)00117-4)

36. Capistrán MA, Moreles MA, Lara B. 2009 Parameter estimation of some epidemic models. The cases of recurrent epidemics caused by respiratory syncytial virus. *Bull. Math. Biol.* 71, 1890–1901. (doi:10.1007/s11538-009-9429-3)

37. Gábor A, Banga JR. 2015 Robust and efficient parameter estimation in dynamic models of biological systems. *BMC Syst. Biol.* 9, 1–25. (doi:10.1186/s12918-015-0219-2)

38. Cheetham N, Waite S, Ebyarimpa I, Leber W, Brennan K, Panovska-Griffiths J. 2021 Determining the level of social distancing necessary to avoid future COVID-19 epidemic waves: a modelling study for North East London. *Sci. Rep.* 11, 5806. (doi:10.1038/s41598-021-84907-1)

39. Teschl G. 2012 *Ordinary differential equations and dynamical systems*. Providence, RI: American Mathematical Society.

40. Zou H. 2008 Fixed point theory and elliptic boundary value problems. In *Handbook of differential equations: stationary partial differential equations*, vol. 6, pp. 503–583. North-Holland, Netherlands: Elsevier. (doi:10.1016/S1874-5733(08)80024-1)

41. Press WH, Vetterling WT, Teukolsky SA, Flannery BP. 1986 *Numerical recipes*. Cambridge, UK: Cambridge University Press.