A comparison of existing measles models

by

Clifford Allotey

A thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Mathematics

University of Manitoba

Winnipeg

Abstract

With improvements in computational power, more and more mechanistic models are being formulated to explain physical and biological phenomena. In this thesis, we assess the strengths and weaknesses of different well-known models of pre-vaccine era measles dynamics. We investigate the assumptions inherent in each model, how well they fit to pre-vaccine era data and how well simple extensions of them perform when extended to modeling vaccine era dynamics. The four focus models we studied were (1) the standard deterministic SEIR models with school-term forcing, (2) the stochastic SEIR model with school-term forcing, (3) the Bjornstad et al. [2002] time series SIR model with a general time-varying transmission rate and (4) the He et al. [2010] model which incorporates school-term forcing, cohort effect and infection from outside the population into the model. We found that the He et al. [2010] model, fitted using maximum likelihood estimation, is the best model in terms of likelihood and the Akaike information criterion. However this model was also the lowest ranked among all four models when comparing fit using residuals, leading to some open questions on possible trade offs between noise and likelihood.

Acknowledgements

I want to thank my supervisor Dr. Felicia Magpantay. I will be forever grateful for her patience, kindness, support, expertise, help and financial assistance throughout my studies. I would also thank the Department of Mathematics for financial support throughout my studies. Last but not the least, I would also thank my mum for the love and support she has given to me.

Contents

List of Figures

List of Tables

1

EPIDEMIOLOGICAL BACKGROUND

1.1   Introduction

Infectious diseases of humans have been well-documented in literature and historical records due to their sometimes calamitous effects on civilizations. Ancient Hindu texts mention the disease of leprosy as early as 2000 BCE. The effect of infectious diseases in shaping the evolution of human society cannot be overstated. Diseases have had repercussions on communities similar to that of warfare and natural disasters. During the 1700s conquest of the Americas, Europeans introduced infections like smallpox that may have devastated the populations of the natives who were ill-equipped to handle such conditions. It is not only humans that have been affected by the courses of epidemics, many other species on earth have been touched and scarred by infectious diseases (Anderson and May [1991], Grenfell and Dobson [1995]). These days we often hear about the Tasmanian devil (*Sarcophilus harrisii*) population becoming endangered in part due to a transmissible facial tumour disease threatening the wild population since the 1990s (Miller et al. [2011]).

The effect of a disease on a population varies, depending on the structure of the population (rural or urban, aging or young, easy or difficult access to health care), and the history that the community has had with the specific disease. In many cases, an individual can develop protection from reinfection after recovery from an infectious disease. This protection may be temporal or permanent. Additionally, immunity may sometimes be induced in an individual by infection with a weaker strain of the infection-causing agent (a virus or bacterium), or by other biological preparations designed to elicit a similar immune response. Mathematical epidemiology is the study of how these individual-level effects translate to ramifications at the population-level, and how public health interventions such as vaccination can change the progression of a disease through a population.

Mathematical and statistical tools have been used to describe the dynamics of infectious diseases since Daniel Bernoulli's smallpox model in 1766. Models are often classified into one of two main categories: (1) empirical models and (2) mechanistic models. Empirical models are based on empirical observations, focusing on collecting data and making predictions from observed data. Though empirical model predictions in the short term can often be accurate, it does not provide additional understanding of the underlying processes involved in a disease process. Mechanistic models, on the other hand, attempt to directly explain the underlying processes involved in a process. These models may not be as good in forecasting since this is not the purpose of such models. To illustrate the difference between empirical and mechanistic models, let us consider a disease that tends to have annual peaks in incidence at a specific time of year. An empirical model can reproduce these dynamics by adopting a form that allows for such peaks, such as a seasonal auto-regressive moving average (SARMA) model. A mechanistic model may include a forcing function that reflects actual seasonally changing circumstances, such as changing transmission rates due to school-term schedules, or annual temperature variation, or the
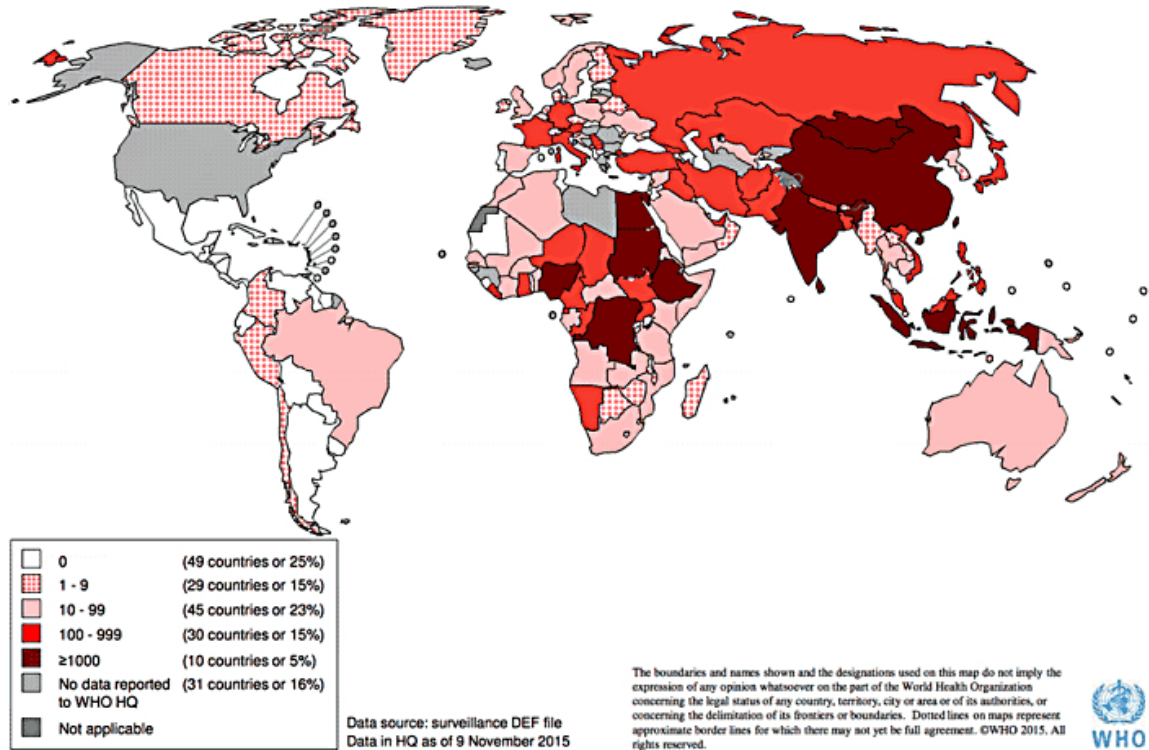
Figure 1.1: Distribution of the number of reported measles cases in the world from April 2015 to Septembers 2015. Image source: WHO 2015.

alternation of wet and dry weather periods.

In this thesis we present a study of different mechanistic models of measles transmission in which assumptions are made about the nature of the underlying disease dynamics based on epidemiological research and medical observations. These models are ranked based on how they perform when their simulations are compared to data. We begin with a short introduction of the disease in section 1.2. In section 1.3 we present some epidemiological terminology that is used in this thesis. In section 1.4 we state the goals of this thesis and in section 1.5 we present an outline of what will be covered in each of the succeeding chapters.

Figure 1.2: A child inflicted by measles displaying measles rashes. Image source: Centers for Disease Control and Prevention (CDC). Photo Credit: Jim Goodson.

1.2 Measles

Measles is an infectious childhood disease caused by the single-stranded RNA virus *Measles morbillivirus* (Barrett [1987]). The mode of transmission can be airborne (via coughs and sneezes from an afflicted person) or via direct contact with tainted surfaces. A person may be infected if the virus comes into contact with the nose, mouth, eyes and skin. There is no vertical transmission which means that newborns may not get the infection from their mothers. However a mother that has had the disease in the past passes temporary immunity to her offspring. This immunity wanes and lasts for approximately four months (Black [1984]).

Measles is now vaccine preventable, however in the pre-vaccine era, most people got infected by age 20 (Anderson and May [1978]), and the mean age of infection was much younger, approximately 4–5 years. After infection, there is a latent stage of

the disease which lasts for approximately 10–12 days (Sume et al. [2012]). After, this stage individuals are usually infectious for approximately 8 days (Centers for Disease Control and Prevention [2015]). According to the Centers for Disease Control [2017], the symptoms of measles include:

- Fever

- Cough

- Runny nose

- Conjunctivities (watery and red eyes)

- Rashes (flat red spots on the skin) appear a few days after the other symptoms

People often recover fully from measles within two weeks, although the disease may last for three weeks. Measles on its own does not usually have disease-related fatalities, however, it can be fatal to infants, people with weakened immune system and malnourished persons (Anderson and May [1991]). There are 134,200 people recorded to have died from measles in 2015 (WHO 2017), which means that the disease is still a leading cause of death in the world, especially for children.

There is no cure for measles, most people recover naturally from it. However, there is a very effective measles vaccine with an efficacy estimated to be between 95% to 98%. This is administered to children at six months or older since maternal immunity transferred to babies can interfere with vaccine efficacy (de Quadros [2004]). However, not everyone can be vaccinated. Exceptions include people who have vaccine allergies, infants too young to be vaccinated, people with HIV and other diseases that weaken the immune system making vaccination life-threatening. These people are susceptible to measles infection and cases of measles sometimes flare up even in developed countries due to recent campaigns by "anti-vaxxers" against vaccination based on the discredited theory that this can cause autism (Hotez [2016]).

In regions where measles vaccine coverage has been high for many years (such as the United States, United Kingdom and Canada), the dynamics of measles can be be roughly split into three different time frames: the pre-vaccine era, the transition era right after the start of mass vaccination, and the vaccine era wherein the disease has settled into stationary dynamics. The dynamics of the disease during the pre-vaccine era is well-studied and many models can adequately reproduce the dynamics of pre-vaccine era measles time series data. However, the same cannot be said about the transient and vaccine eras (Barrett [1987], Grenfell et al. [2014]). In this thesis we reviewed and compared existing measles models for the pre-vaccine era. We performed a preliminary comparison of how these models perform when extended into the vaccine era. Lastly, we also presented suggestions to improve the fit to the transient and vaccine eras.

1.3   Terminology

We begin by introducing some of the standard terminology used in the mathematical modeling of disease transmission. The main references for this section are Milwid et al. [2016] and Mishra et al. [2010].

- **Host:** This is the living entity of interest that can be infected by a pathogen (Mishra et al. [2010]). In this thesis, we focus on humans as the hosts for measles.

- **Pathogen:** This is the organism that infects the host (Mishra et al. [2010]). In this thesis the pathogen we consider is the measles virus.

- **Endemic:** This refers to the state of constant presence and/or usual prevalence of a disease or infectious agent in a population within a geographic area (Dicker [2006]).

- **Epidemic:** There is said to an epidemic of a disease if there is a substantial increase in incidence of the disease to a peak (the epidemic peak), then a substantial decrease to pre-epidemic levels.

- **Compartmental models:** These are systems of equations describing flows of material between units called compartments (Jacquez and Simon [1993]). These compartments partition the population under study at any given time. For each compartment, its elements are time-homogeneous (local homogeneity). In the epidemiological setting, these are models that group hosts, based on their states at infection (Mishra et al. [2010]).

- **SEIR model framework:** This is a standard compartmental model framework used in epidemiology in which the hosts have the following compartments available: susceptible, exposed/latent, infectious/infected, recovered/removed. These terms are described below and further details are provided in Chapter 2. An SIR model framework is like the SEIR model framework but without the exposed/latent stage.

- **Susceptible:** This is the compartment of individuals who can be infected by the pathogen.

- **Exposed/Latent:** This is the compartment of individuals that have been infected by the pathogen but are still in the latent stage of the disease and are not yet able to transmit the disease at that particular time. Additionally, the latent period is the time in between successful infection of the host and its disease transmission stage. In this thesis we usually refer to this compartment as simply the exposed compartment.

- **Infectious/Infected:** This is the compartment of hosts that have been infected by the disease and is also infectious, i.e. capable of transmitting the

disease to susceptible hosts. In this thesis we usually refer to this compartment as simply the infectious compartment.

- **Recovered/Removed:** In this thesis, we consider the recovered compartment to be the same as the removed compartment. However some other papers may make a distinction between these compartments. The recovered compartment is usually defined to be the compartment of individuals that have recovered from the infection and is no longer infectious. The removed compartment is usually defined to be the compartment of individuals that are not currently exposed or infected, and are not susceptible to the disease. If this distinction is made then hosts that have been vaccinated (with a perfect vaccine) are in the removed compartment but not in the recovered compartment. In this thesis we assume that all recoveries and vaccination lead to the recovered/removed compartment and simply refer to this compartment as the removed compartment. Hosts in this compartment do not take part in the the transmission process.

- **Incidence:** This is the number of new cases of the disease in a span of time, usually normalized by the size of the population (usually in units of cases per 100,000 individuals).

- **Prevalence:** Prevalence is the number of cases of the disease at a particular point in time. This is equal to the size of infected classes at the certain point in time.

- **Vaccination:** Vaccination is the act of inducing immunity in a host, by exposing the host to a weaker strain of the pathogen or by a biological preparation that is designed to elicit an immune response similar to that of natural infection (Centers for Disease Control and Prevention [2015]).

- **Immunity:** Immunity as defined in Milwid et al. [2016], is protection against

a disease. Individual immunity (whether temporal or permanent) can be acquired at birth from maternal antibodies, recovery from a disease, or vaccination. Herd immunity is a population-level protection, obtained as a result of collective individual immunity (Milwid et al. [2016]).

- **Basic reproduction number** R_0**:** This is the average number of secondary infections from the introduction of an infected host into a completely susceptible population. For deterministic models, this is the spectral radius of the next generation matrix associated with the transmission model. The technical definition is given in Chapter 2.

- **Deterministic model:** A deterministic model is a system in which no randomness is involved in determining the states of the system. For example, compartmental models written as a set of ordinary differential equations (ODEs) or partial differential equations (PDEs) are deterministic models. Deterministic models are usually written as models that reflect the "average" behaviour of the system (Mishra et al. [2010]). In this thesis what we refer to as deterministic models can still allow for randomness in the measurement/observations of the states, but not in the states themselves. For example, one can define a deterministic SEIR model defined by ODEs that yields a fixed number of states within some time interval. However the actual observation process (reporting of the cases to authorities to be recorded as data) may involve a reporting probability and a binomial distribution in reports.

- **Stochastic model:** A stochastic model is described by processes that incorporate some randomness in determining the states of the system (Mishra et al. [2010]). Each transition denotes an event that can occur to each individual in a time interval according to a probability that is determined using rates of transitions. A stochastic model needs to have stochastic processes deter-

mining the states of the model, and it may also allow for stochasticity in the measurement/observation of the states.

- **Trajectory:** A trajectory is a solution of a deterministic process. In this thesis we use trajectory to mean the states of a deterministic model as a function of time. These are fully determined for a given deterministic model with fixed parameter values and initial conditions.

- **Simulation:** A simulation is a realization of a stochastic process. There can be many different realizations of the same model. One can also present a simulation of the stochastic observations of a deterministic process.

- **Generation time:** The generation time of a disease like measles is defined to be the time interval from the start of the infectious period of a primary case of infection to the beginning of the infectious period in a secondary case.

1.4   Goals of this project

The transmission dynamics of measles has been widely studied and there are many models that have been developed to explain its dynamics, most of them focusing on the pre-vaccine era. However there are still some inconsistencies in the best models we have of measles. Some of the best models fitted to time series of measles reports yield estimates of the infectious period of measles that is known to be different from clinical observation. Others yield estimates of the basic reproduction number of measles that are much higher than would be expected from serological data. In this thesis, we reviewed and compared four well-known models of pre-vaccine era measles transmission that are based on the standard susceptible-exposed-infected-recovered (SEIR) model framework:

1. Deterministic SEIR model with school-term forcing

2. Stochastic SEIR model with school-term forcing

3. Bjornstad et al. [2002] time series SIR (TSIR) model with a general forcing function

4. He et al. [2010] stochastic SEIR model with school-term forcing, the cohort effect and infection from outside the population

We recreated the existing analysis, coded and refitted each model to bi-weekly data from the pre-vaccine era in London, United Kingdom (UK). We evaluated the performance of these models using likelihood and other model comparison criteria such as Akaike information criterion (AIC) and residuals. Additionally, to compare the performance of these models into the transient and vaccine eras, we compared the simulations of these models using the pre-vaccine era estimated parameters and data on UK national vaccine uptake coverage. At the end of this thesis we present a discussion of the performance of each model, and a discussion of what these model comparison criteria mean, and some suggestions on how to improve model fit in the vaccine era.

1.5 Outline of the thesis

In Chapter 2, we present the mathematical preliminaries required for this study. We review models such as the standard ODE form of the SEIR model (without any forcing), the SEIR model with the addition of annual forcing functions, and the continuous time Markov chain SEIR model. We also review the partially observed Markov process model (POMP) which is the set up we use to make a clear distinction between state and observation processes. We discuss how to calculate the likelihood of a POMP model and how this can be calculated using the particle filter. Finally, we present the Akaike information criterion (AIC) which is a criterion for model

comparison. In Chapter 3 we present a review of some existing measles models, including the four measles model that we focused on in this thesis. We discuss each focus model in detail, outlining the assumptions involved in the formulation of the model, its strengths and weaknesses, and the difficulties in simulating and fitting the parameters of the different models. In Chapter 4 we compare our four focus models using standard comparison criteria, such as the AIC and residuals. We also discuss the features apparent in simulations of the models and compare it to the data. In Chapter 5 we present a summary of the thesis and discuss some future directions for modeling measles past the pre-vaccine era.

2

MATHEMATICAL
BACKGROUND

In this chapter, we present the mathematical preliminaries for this study. In section 2.1, we present the standard ODE formulation of the susceptible-exposed-infected-recovered (SEIR) model with demography. In section 2.1.1 we go through basic properties and analysis of this model. In section 2.1.2, we explain why this autonomous model is inadequate to describe the dynamics of measles. In section 2.2, we expand upon the standard SEIR model by introducing a seasonal forcing function to allow for periodic solutions, which can look like either the annual and biennial dynamics of measles in pre-vaccine era London, UK (shown in Figure 2.1). In section 2.3, we present the steps to derive a stochastic SEIR model from the deterministic models. In section 2.5 the definition of the likelihood is given and the expressions for the likelihood for both deterministic and stochastic models are derived. Maximum likelihood estimation of parameters of the model is also discussed. In section 2.6, we present the particle filter, a Monte Carlo method for calculating the likelihoods. In section 2.7, we discuss the Akaike Information Criterion (AIC), a criterion for selecting the best among competing models.
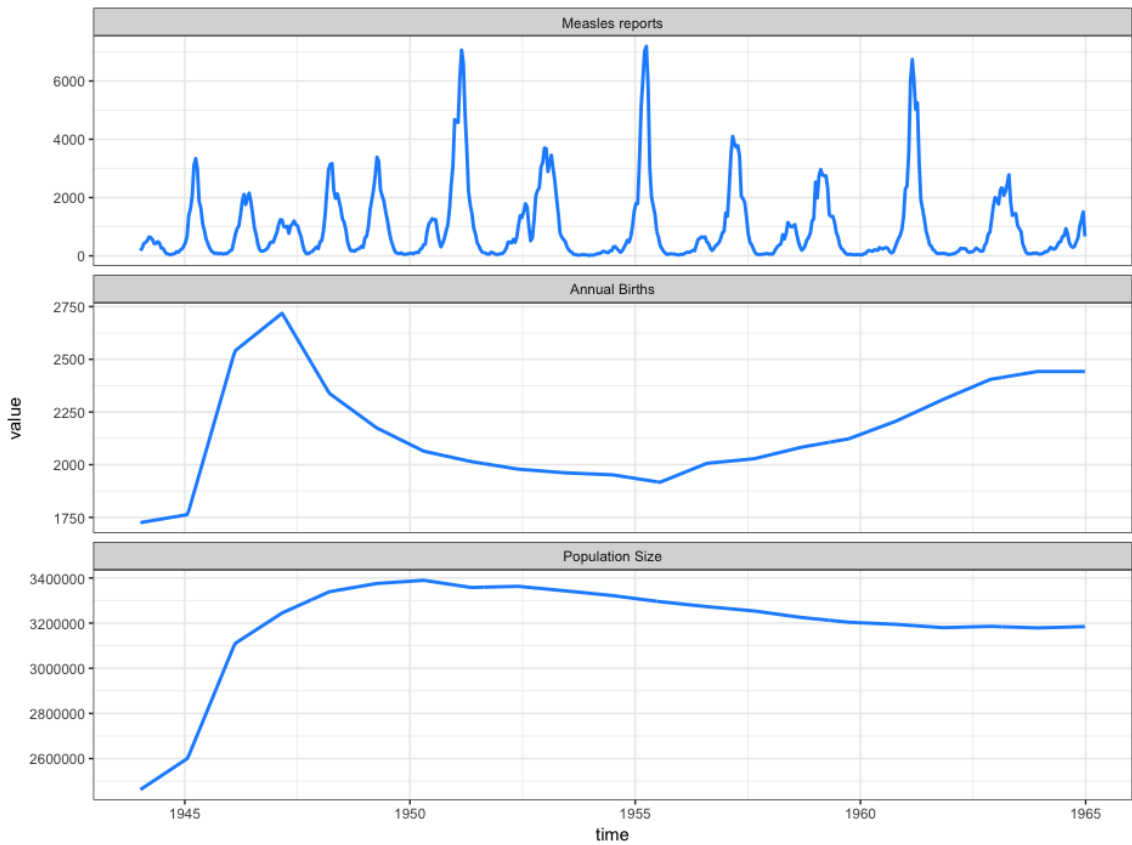
Figure 2.1: A plot of reported measles cases, the annual birth rates and population sizes of London from 1944 to 1965. Source of data: `http://kingaa.github.io/pomp/vignettes/twentycities.rda`

2.1 Standard SEIR model

The SEIR model is an extension of the SIR model first studied by Kermack and McKendrick [1932, 1933]. SEIR and SIR models have now been studied extensively in the literature (Hethcote [2000, 1976]). In this section we review the basic assumptions and properties of the standard SEIR model.

We begin by considering a population with birth rate of $\mu_b > 0$ and per capita death rate of $\mu > 0$. Thus the total population $N(t)$ satisfies

$$N'(t) = \mu_b - \mu N(t),$$

which converges to its state value $N_0 = \frac{\mu_b}{\mu}$. For convenience we assume that $N(0) = N_0$ and this implies that $N(t) = N_0$ for all t. We now divide the total population into the SEIR components and assume that we always have,

$$S(t) + E(t) + I(t) + R(t) = N(t) = \frac{\mu_b}{\mu}.$$

Let $\beta > 0$ denote the transmission rate of the disease through this population, $\sigma > 0$ be the incubation rate (the rate of going from the latent to infectious stage of the disease), and $\gamma > 0$ be the recovery rate (the rate of leaving the infectious period). Let $v \in [0, 1)$ be the vaccination coverage (which is equal to zero in the pre-vaccine era). We assume that the disease-induced death rate is negligible in this model so that the constant death rate μ applies to all the compartments. We also assume that there is no vertical transmissions, i.e. unvaccinated newborns are born directly into the susceptible class. Under the assumption of homogeneous populations mass action, there are $\beta S I / N_0$ infections per unit time, going from the susceptible compartment to the exposed compartment. Individuals in the exposed compartment move on to the infected compartment with rate σ, then eventually into the recovered

class with rate γ. We assume that both vaccinated and recovered individuals have permanent immunity and thus the fraction v of newborns that have been vaccinated directly goes into the recovered class. The system of equations in (2.1)–(2.4) and a schematic of this model is presented in Figure 2.2.

$$S' = \mu_b(1-v) - \beta IS/N - \mu S \tag{2.1}$$

$$E' = \beta SI/N - (\sigma + \mu)E \tag{2.2}$$

$$I' = \sigma E - (\gamma + \mu)I \tag{2.3}$$

$$R' = \gamma I - \mu R + \mu_b v \tag{2.4}$$
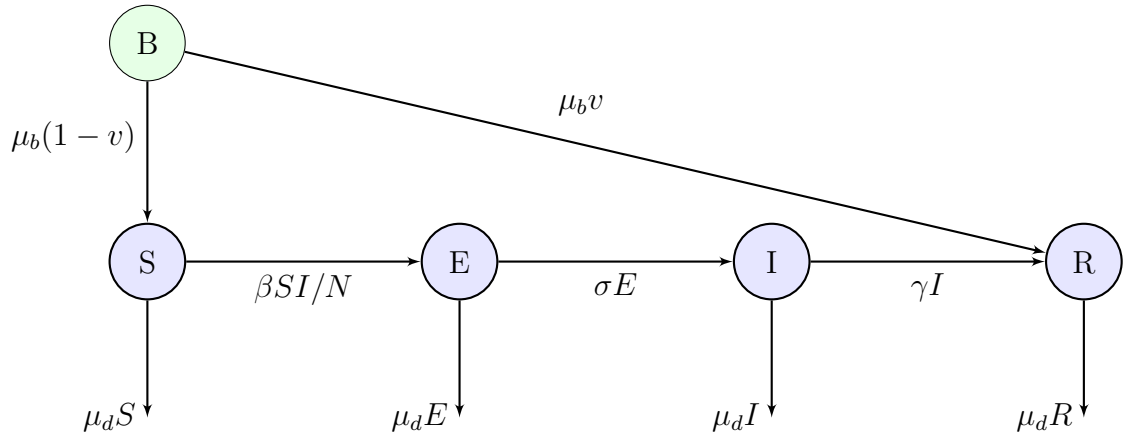
Figure 2.2: Diagram of the SEIR model

We require the following assumptions on the initial conditions of the system,

$$S(0) \geq 0, \quad E(0) \geq 0, \quad I(0) \geq 0, \quad R(0) \geq 0 \tag{2.5}$$

and

$$S(0) + E(0) + I(0) + R(0) = \frac{\mu_b}{\mu}. \tag{2.6}$$

Before we evaluate how suitable this model is for modeling measles dynamics, we first ensure that the model is well-posed and it at least has a solution.

16

2.1.1 Properties

Invariance

To ensure that the model is well-posed, we have to verify that the states, which represent populations, cannot be negative given non-negative initial conditions. Thus we expect that solutions of the equations with non-negative initial condition remain in the following simplex:

$$T = \left\{(S, E, I, R) \mid S \geq 0, E \geq 0, I \geq 0, R \geq 0, \ S + E + I + R = \frac{\mu_b}{\mu}\right\}.$$

We have already shown that $N = S + E + I + R$ has a steady sate at $\frac{\mu_b}{\mu}$ if the requirements on initial conditions are satisfied. To prove that T is invariant, let us examine the behaviour of states at the boundaries of this domain.

1. At the boundary $S = 0$ the equation for S' becomes $S' = \mu_b(1 - v) > 0$. Thus the solution cannot exit T by crossing this boundary.

2. At the boundary $E = 0$ the equation for E' becomes $E' = \beta S(t)I(t)/N \geq 0$. If $E(t) = 0$, $S(t) > 0$ and $I(t) > 0$ then $\beta S(t)I(t)/N > 0$ and the solution cannot exit T by crossing the $E = 0$ boundary in this case. If $E(t) = 0$, $S(t) = 0$ and $I(t) > 0$ we have $E' = 0$ and $S' > 0$, thus $E'' > 0$ and the solution cannot cross the $E = 0$ boundary. If $E(t) = 0$, $I(t) = 0$ and $S(t) \geq 0$ then the SEIR equations become:

$$S' = \mu_b(1 - v) - \mu S \tag{2.7}$$

$$E' = 0 \tag{2.8}$$

$$I' = 0 \tag{2.9}$$

$$R' = -\mu R + \mu_b v \tag{2.10}$$

Thus in this case the solution maintains $E(s) = I(s) = 0$ for all $s > t$ and additionally, $S(s) = S(t)e^{-\mu(s-t)} + \frac{\mu_b(1-v)}{\mu}((1 - e^{-\mu(s-t)})$ and $R(s) = R(t)e^{-\mu(s-t)} + \frac{\mu_b v}{\mu}((1-e^{-\mu(s-t)})$ which keeps the solution in T. Thus the solution cannot exit T by crossing the boundary $E = 0$.

3. At the boundary $I = 0$ the equation for I' becomes $I' = \sigma E$. If $I(t) = 0$ and $E(t) > 0$ then $I'(t) = \sigma E(t) > 0$ and thus the solution cannot exit T through the $I = 0$ boundary in this case. The case $I(t) = 0$ and $E(t) = 0$ has already been considered above. Thus the solution cannot exit T via the $I = 0$ boundary.

4. At the boundary $R = 0$ the equation for R' becomes $R' = \gamma I + \mu_b v \geq 0$. If $R(t) = 0$ and $I(t) > 0$ then $R'(t) = \gamma I(t) + \mu_b v > 0$ and the solution cannot exit T via this boundary. In the case when $R(t) = 0$ and $I(t) = 0$, we consider cases: First, the case when $R(t) = I(t) = E(t) = 0$ is included in the discussion of the case $I(t) = E(t) = 0$ above and we know the the solution cannot leave T in this case. Second, the case when $R(t) = I(t) = 0$ and $E(t) > 0$ yields $R'(t) = \mu_b v \geq 0$. If $v > 0$ then $R'(t) > 0$ and the solution cannot exit T by crossing the $R = 0$ boundary. If $v = 0$ then $R'(t) = 0$ and $R''(t) = \gamma I'(t) = \sigma E(t) > 0$ which also shows that the solution cannot exit T by crossing the $R = 0$ boundary.

This completes the proof of the invariance of the domain T.

Existence and uniqueness of solutions

Let the states $X = (S, E, I, R)$ and $f(X)$ be given by,

$$f(X) = \begin{bmatrix} \mu_b(1-v) - \beta SI/N - \mu S \\ \beta SI/N - (\sigma + \mu)E \\ \sigma E - (\gamma + \mu)I \\ \gamma I - \mu R + \mu_b v \end{bmatrix}$$

Since f is continuous in X and $f'(X)$ is continuous at all points in its domain then f is Lipschitz. From the standard theory of ordinary differential equations, this proves local existence and uniqueness of solutions (Miller and Michel [1982]). Since we also have X confined to T, solutions are always bounded yielding existence and uniqueness in its domain (i.e. the simplex T)

Disease-free equilibrium

A point X is an equilibrium point of the system if $f(X) = 0$. The disease-free equilibrium $X^0 = (S^0, E^0, I^0, R^0)$ of the model is an equilibrium with $I^0 = 0$. Thus the disease-free equilibrium of the system we have defined can be found by setting $I^0 = 0$ in the equation $f(X^0) = 0$ and solving for S^0, E^0 and R^0. Clearly this equilibrium also requires $E^0 = 0$ and we can solve for the remaining terms easily. We obtain the following expression for the unique disease-free equilibrium of this system,

$$X^0 = \Big(\frac{\mu_b(1-v)}{\mu}, 0, 0, \frac{\mu_b v}{\mu} \Big). \tag{2.11}$$

Basic reproduction number

The basic reproduction number (denoted by R_0) is defined to be the spectral radius of the the next generation matrix (Diekmann et al. [2010],van den Driessche and Watmough [2008]). To derive this we consider the states with infection, which are

E and I.

$$E' = \beta SI/N - (\sigma + \mu)E \tag{2.12}$$

$$I' = \sigma E - (\gamma + \mu)I. \tag{2.13}$$

We linearize (2.12)–(2.13) about the disease-free equilibrium and decompose the linearized system into,

$$\begin{bmatrix} E \\ I \end{bmatrix}' = \left(T + \Sigma\right) \begin{bmatrix} E \\ I \end{bmatrix},$$

where

$$T = \begin{bmatrix} 0 & \beta \frac{\mu_b(1-v)}{\mu} \\ 0 & 0 \end{bmatrix}, \quad \Sigma = - \begin{bmatrix} \sigma + \mu & 0 \\ -\sigma & \gamma + \mu \end{bmatrix}.$$

Here T is the matrix that include the terms giving rise to new infections, and Σ is the matrix taking into account the progression of the disease. The next generation matrix is $K = -T\Sigma^{-1}$ evaluated at the disease-free equilibrium. Thus,

$$K = \begin{bmatrix} \frac{\mu_b(1-v)}{\mu} \frac{\beta}{\gamma+\mu} \frac{\sigma}{\sigma+\mu} & \frac{\mu_b(1-v)}{\mu} \frac{\beta}{\gamma+\mu} \\ 0 & 0 \end{bmatrix}$$

Thus, $R_0 = \frac{\mu_b(1-v)}{\mu} \frac{\beta}{\gamma+\mu} \frac{\sigma}{\sigma+\mu}$. Theory by Diekmann et al. [2010] shows that if $R_0 < 1$, the disease-free equilibrium is locally stable. We can break down the terms involved in the basic reproduction number in the following manner:

$$R_0 = \underbrace{\frac{\mu_b(1-v)}{\mu}}_{\substack{\text{susceptible} \\ \text{population}}} \underbrace{\beta}_{\substack{\text{transmission} \\ \text{rate}}} \underbrace{\frac{1}{\gamma+\mu}}_{\substack{\text{average} \\ \text{infectious} \\ \text{duration}}} \underbrace{\frac{\sigma}{\sigma+\mu}}_{\substack{\text{probability} \\ \text{progressing} \\ \text{from E to I}}} \tag{2.14}$$

Endemic equilibrium

The endemic equilibrium of the SEIR model is the unique equilibrium with postive number of infections $I > 0$. We can denote this by $X^* = (S^*, E^*, I^*, R^*)$. We can derive expressions for this by first finding the S^*, E^* and R^* components in terms of I^*. By setting the right hand sides of (2.1), (2.2) and (2.4) equal to zero we derive,

$$S^* = \frac{\mu_b(1-v)}{\mu + \frac{\beta I^*}{N}} \tag{2.15}$$

$$E^* = \frac{\mu_b(1-v)}{\mu + \frac{\beta I^*}{N}} \frac{\beta I^*}{N(\sigma + \mu)} \tag{2.16}$$

$$R^* = \frac{\gamma I^* + \mu_b v}{\mu} \tag{2.17}$$

An expression for I^* can be found by deriving another expression for E^* using (2.3), setting this equal to the E^* value derived above and solving for I^*.

$$\frac{\gamma + \mu}{\sigma} = \frac{\mu_b(1-v)}{\mu + \frac{\beta I^*}{N}} \frac{\beta}{N(\sigma + \mu)} \tag{2.18}$$

$$1 = \frac{\mu}{\mu N + \beta I^*} \frac{\mu_b(1-v)}{\mu} \frac{\beta}{\gamma + \mu} \frac{\sigma}{\sigma + \mu} \tag{2.19}$$

$$1 = \frac{\mu}{\mu_b + \beta I^*} R_0 \tag{2.20}$$

$$I^* = \frac{1}{\beta}(\mu R_0 - \mu_b) \tag{2.21}$$

From the results of Li and Muldowney [1995], this endemic equilibrium is globally asymptotically stable in the simplex T when $R_0 > 1$.

2.1.2 Shortcomings of the standard SEIR model

The fact that the endemic equilibrium X^* is globally asymptotically stable in the simplex T when $R_0 > 1$ means that solutions of (2.1) must all asymptotically tend to the endemic equilibrium point X^*. Thus the standard SEIR model we have reviewed (with constant birth rates and transmission rates) cannot reproduce the sustained
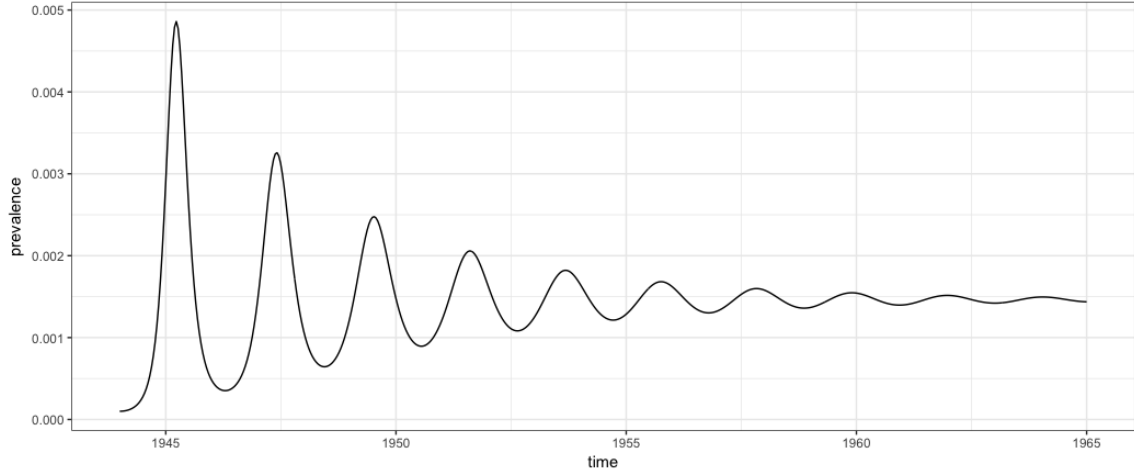
Figure 2.3: A sample trajectory of the proportion of prevalence $(E(t)/N(t) + I(t)/N(t))$ resulting from the standard SEIR model. When $R_0 > 1$ the model is known to asymptotically tend to an endemic equilibrium value and its solutions does not include sustained oscillations of epidemics.

periodic dynamics such as the annual and biennial dynamics of the disease that have often been observed in pre-vaccine era data (e.g. pre-vaccine era data in London shown in Figure 2.1). Even if the model is changed to include changing population sizes and birth rates (accomplished by making μ_b and μ functions of time), as long as these are not strongly periodic forcing functions, the model at best can exhibit transient oscillations that eventually get damped. A typical trajectory of this model is shown in Figure 2.3.

2.2 SEIR model with school-term forcing

Table 2.1: The start and end dates of each of the four school terms in the UK. This is used in the school-term forcing function.

School Session	Session Start Date	Session End Date
1	7th January	10th April
2	15th April	18th July
3	9th Sept	27th Oct
4	4th Nov	31 Dec

One obvious way to improve the standard SEIR model to enable it to exhibit the periodic dynamics seen in the data is to incorporate a periodic forcing function such as a time-dependent transmission rate (Earn et al. [2000a]). Such a forcing function can be consistent with real phenomena. In this section we introduce a forcing function that allows for variation in transmission rate based on the schedule of school-terms. Measles is a childhood disease, and the amount of contacts between children is expected to be significantly higher during the school term and lower during school breaks. Thus it makes sense to incorporate this into the model. This can be done by changing the system of differential equations from (2.1)–(2.4) into one with a school-term forcing function:

$$S' = \mu_b(1-v) - B(t)\frac{SI}{N} - \mu S \tag{2.22}$$

$$E' = B(t)\frac{SI}{N} - (\lambda + \mu)E \tag{2.23}$$

$$I' = \lambda E - (\gamma + \mu)I \tag{2.24}$$

$$R' = \gamma I - \mu R + \mu_b v \tag{2.25}$$

The function $B(t)$ is a piece-wise continuous function defined to be,

$$B(t) = \begin{cases} \beta\,(1 + 2(1-p)a), & \text{during term}, \\ \beta\,(1 - 2pa), & \text{during vacation}, \end{cases} \tag{2.26}$$

where a is the amplitude parameter, and p the proportion of the school days for which the school-term is in session. Table 2.1 shows the start and end dates of days for which the school-term is in session in the UK and this what we use in the school-term forcing function. All parameters above are non negative. We also require the same assumptions on the initial conditions (2.5)–(2.6). Figure 2.4 shows a trajectory of the SEIR model with forcing.
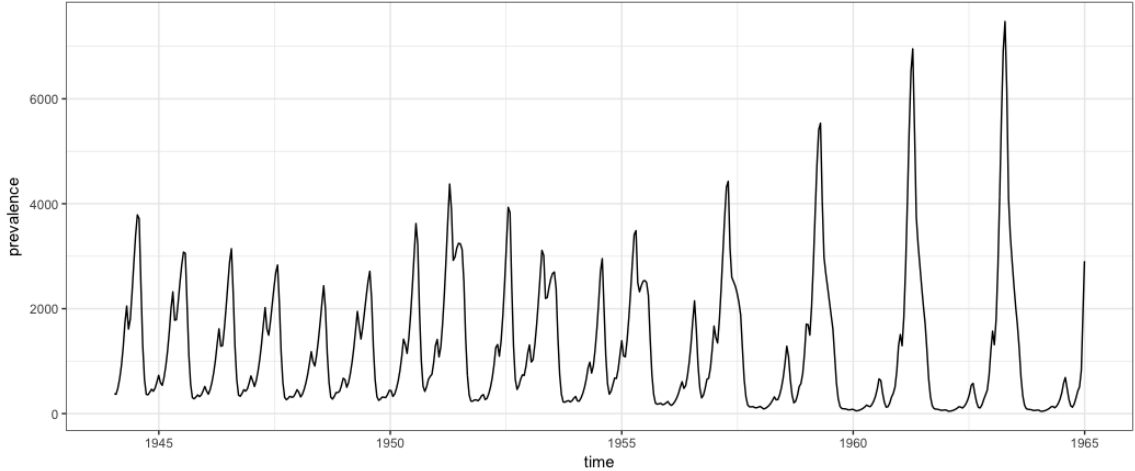
Figure 2.4: A sample trajectory of the number of cases from an SEIR model with school-term forcing. Unlike the standard SEIR model with no forcing, this model has periodic endemic solutions.

The invariance of the simplex T can be proven in a similar manner for the SEIR model with seasonal forcing. The disease-free equilibrium can be determined by setting $E = I = 0$ and obtaining the same disease-free equilibria as in (2.11). We also no longer have an endemic equilibrium point, but rather an endemic periodic solution. The stability analysis of the zero solution and the endemic periodic solution requires using the method of Floquet Multipliers (Williams and Dye [1997]) which is beyond the scope of this thesis.

Both the standard SEIR model and the SEIR model with school-term forcing as presented here are deterministic processes. Next we present models of epidemics that allow for some stochasticity in the dynamics.

2.3 Stochastic SEIR model

Markov chains can be used to model epidemic dynamics. A system is said to have the Markov property if the current state of the system only depends on of its previous state. Mathematically this property can be stated in the following manner: Let $X(t)$ be a stochastic process. This is a Markov process if, given the value of $X(t)$

the value of $X(t + h)$ for some $h > 0$ is independent of $X(s)$ for all $s < t$. The Markov property lets us specify a model by giving the transition probabilities on small intervals of time. This Markov property is inherent in many models when all relevant variables are included in the states and the SEIR models can be conveniently remodeled as counting processes (He et al. [2010]). Following Bretó et al. [2009] and He et al. [2010], in this section we present a derivation of a stochastic version of the SEIR model by tracking the flow of individuals between each compartment to keep track of the number of individuals in each compartment.

The stochastic SEIR model that we derive here is a continuous-time, discrete state Markov process which can take on values from $\mathbb{N}^4$ (the set of vectors with four components each taking on values in the set of non-negative integers). We let the state variable be denoted by $X(t) = (S(t), E(t), I(t), R(t))$, and the components represent the numbers of individuals at time t in each of the compartments S, E, I and R. The process is defined in terms of counting processes which are non-negative, non-decreasing, integer-valued processes. Let $Z_{A \to B}(t)$ be the total number of individuals that have transitioned from some compartment A to another compartment B in the time interval $[0, t]$. We also use the notation $Z_{O \to A}(t)$ to represent the total number of entry into compartment A from outside the existing population (in this model we assume this is via birth) and $Z_{A \to O}(t)$ to represent exit from compartment A (in this model we assume this is via death). To define the stochastic model, we first specify the probability of a change in $Z_{A \to B}(t)$ after a small time interval, for all pairs of compartments.

Let $t > 0$ and $h > 0$. We define $\Delta Z_{A \to B}(t, t + h) = Z_{A \to B}(t + h) - Z_{A \to B}(t)$ to be the change in the $Z_{A \to B}$ counting process over the time interval $[t, t + h]$. As usual

25

Let $N(t) = S(t) + E(t) + I(t) + R(t)$ and,

$$\mathbb{P}\left(\Delta Z_{S \to E}(t, t+h) = 1 \mid X(t)\right) = \frac{\beta S(t) I(t)}{N(t)} h + o(h) \tag{2.27}$$

$$\mathbb{P}\left(\Delta Z_{E \to I}(t, t+h) = 1 \mid X(t)\right) = \sigma E(t) h + o(h) \tag{2.28}$$

$$\mathbb{P}\left(\Delta Z_{I \to R}(t, t+h) = 1 \mid X(t)\right) = \gamma I(t) h + o(h) \tag{2.29}$$

where $o(h)$ stands for the standard little o notation. We also consider the probability of births into the susceptible class denoted by $\mathbb{P}\left(\Delta Z_{O \to S}(t, t+h) = 1 \mid X(t)\right)$, probability of births into the recovered class (due to vaccination) denoted by $\mathbb{P}\left(\Delta Z_{O \to R}(t, t+h) = 1 \mid X(t)\right)$ and probability of deaths from each of the S, E, I and R components denoted by $\mathbb{P}\left(\Delta Z_{S \to O}(t, t+h) = 1 \mid X(t)\right)$, $\mathbb{P}\left(\Delta Z_{E \to O}(t, t+h) = 1 \mid X(t)\right)$, $\mathbb{P}\left(\Delta Z_{I \to O}(t, t+h) = 1 \mid X(t)\right)$ and $\mathbb{P}\left(\Delta Z_{R \to O}(t, t+h) = 1 \mid X(t)\right)$ respectively,

$$\mathbb{P}\left(\Delta Z_{O \to S}(t, t+h) = 1 \mid X(t)\right) = \mu_b (1 - v) h + o(h) \tag{2.30}$$

$$\mathbb{P}\left(\Delta Z_{O \to R}(t, t+h) = 1 \mid X(t)\right) = \mu_b v h + o(h) \tag{2.31}$$

$$\mathbb{P}\left(\Delta Z_{S \to O}(t, t+h) = 1 \mid X(t)\right) = \mu S(t) h + o(h) \tag{2.32}$$

$$\mathbb{P}\left(\Delta Z_{E \to O}(t, t+h) = 1 \mid X(t)\right) = \mu E(t) h + o(h) \tag{2.33}$$

$$\mathbb{P}\left(\Delta Z_{I \to O}(t, t+h) = 1 \mid X(t)\right) = \mu I(t) h + o(h) \tag{2.34}$$

$$\mathbb{P}\left(\Delta Z_{R \to O}(t, t+h) = 1 \mid X(t)\right) = \mu R(t) h + o(h) \tag{2.35}$$

We assume that the probability of multiple transitions occurring within the time interval $[t, t+h]$ are $o(h)$. The changes in the sizes of the compartments and subcom-

partments over the time interval $[t, t+h]$ interval is given by the following equations.

$$\Delta S = \Delta Z_{O \to S}(t, t+h) - \Delta Z_{S \to E}(t, t+h) - \Delta Z_{S \to O}(t, t+h) \tag{2.36}$$

$$\Delta E = \Delta Z_{S \to E}(t, t+h) - \Delta Z_{E \to I}(t, t+h) - \Delta Z_{E \to O}(t, t+h) \tag{2.37}$$

$$\Delta I = \Delta Z_{E \to I}(t, t+h) - \Delta Z_{I \to R}(t, t+h) - \Delta Z_{I \to O}(t, t+h) \tag{2.38}$$

$$\Delta R = \Delta Z_{O \to R}(t, t+h) + \Delta Z_{I \to R}(t, t+h) - \Delta Z_{R \to O}(t, t+h) \tag{2.39}$$

2.4 Partially observed Markov process model

Partially observed Markov process (POMP) models are more frequently known as hidden Markov models or state-space models. These models consist of unobserved states $X(t)$ and observed states $Y(t)$. For instance, in the disease models that we have been considering, the $S(t)$, $E(t)$, $I(t)$ and $R(t)$ states are all unobserved states so we can as before set $X(t) = (S(t), E(t), I(t), R(t))$ We can also introduce a function $C(t, t+h)$ based on these states which gives the number of new cases of disease generated within a given time interval $[t, t+h]$. Some of these cases will be reported and the number of reported cases would be the observed state $Y(t)$. These reported cases would be the state that should be compared to disease reporting data when models are fitted.

A POMP model requires a distribution of the initial states, a process model describing the transitions of the unobserved states and an observation model describing how observations are taken from the underlying states (King [2016b]) The process model is assumed to satisfy the Markov property. We also assume that the observation process is defined by a distribution $\mathbb{P}(Y(t) = y | X(t) = x(t))$. Thus we can think of the data, $\{y_i^*\}$ as being a single realization of the $\{Y(t)\}$ process.

The general form of a POMP model presented above is a stochastic model with a stochastic process model and a stochastic measurement model. It is also possible

to define a deterministic POMP model by choosing a deterministic process model with a stochastic measurement model. There are also other combinations possible (such as a stochastic process model with deterministic observation) however they are beyond the scope of this thesis.

2.5   Likelihood

Let $y_{1:n}^*$ be a sequence of n observed data at discrete time points $t_{1:n}$. Let $\theta \in \Theta \subset \mathbb{R}^K$ be a vector of parameter values for a given statistical model. The likelihood $\mathcal{L}(\theta)$ as a function of θ is defined to be,

$$\mathcal{L}(\theta) = \mathbb{P}(Y(t_{1:n} = y_{1:n}^*|\theta)), \tag{2.40}$$

(Fisher [1922]). This is a measure of the plausibility that an observed data comes from a distribution given by the set of parameters θ. For a continuous state process, one may also consider the likelihood function by replacing the probability with the probability density function.

For this section and the remaining sections of this chapter we use the notation $X_i = X(t_i)$ and $Y_i = Y(t_i)$. Given a deterministic POMP model, the likelihood has a simple form because for any given parameter θ, the states X_i have only one possible value $x_i(\theta)$ for $i = 1, \ldots, n$. Thus

$$\mathcal{L}(\theta) \;\; = \;\; \mathbb{P}(Y_{1:n} = y_{1:n}^*; \theta) \tag{2.41}$$

$$= \;\; \prod_{i=1}^{n} P(Y_i = y_i^*|X_i = x_i(\theta)) \tag{2.42}$$

It is common to consider the logarithm of the likelihood instead of the likelihood $\mathcal{L}(\theta)$ to convert products into sums in the likelihood expressions like the one above.

We define the log-likelihood to be,

$$\ell(\theta) = \log\left(\mathcal{L}(\theta)\right) \tag{2.43}$$

For the deterministic POMP model, we easily derive that,

$$\ell(\theta) = \log\prod_{i=1}^{n} P(Y_i = y_i^*|X_i = x_i(\theta)), \tag{2.44}$$

$$= \sum_{i=1}^{n}\log P(Y_i = y_i^*|X_i = x_i(\theta)). \tag{2.45}$$

The likelihood of a stochastic POMP model is more complicated since the states are not fixed (for a given θ there are many possible values of X_i). Thus the likelihood needs to be summed over all possible values of the unobserved states according to their transition probabilities,

$$\mathcal{L}(\theta) = \mathbb{P}(y_{1:n}^*|\theta) = \sum_{x_1}\sum_{x_2}\cdots\sum_{x_n}\prod_{i=1}^{n}\mathbb{P}(Y_i = y_i^*|X_i = x_i, \theta)\mathbb{P}(x_i|x_{i-1}, \theta). \tag{2.46}$$

The log-likelihood $\ell(\theta)$ of a stochastic POMP model is simply the logarithm of the expression above.

The maximum likelihood estimate (MLE) is defined to be the value of the parameter vector θ that yields the highest value of the likelihood, or equivalently the log-likelihood.

$$\theta_{\text{MLE}} = \arg\max_{\theta\in\Theta}\mathcal{L}(\theta) = \arg\max_{\theta\in\Theta}\ell(\theta).$$

This may not be unique.

2.6 Particle filter

The expression (2.46) is in general not analytically tractable. However, it can be estimated by Monte Carlo methods. For instance, this can be estimated using the

standard Monte Carlo method by just generating multiple simulations of the model, finding the likelihood of the data for each simulation and averaging across all simulations. Though this is known to converge to the true likelihood as the number of simulations approach infinity, it can require an unfeasibly large amount of simulations to converge (King [2016b]). Instead of using standard Monte Carlo methods to estimate (2.46) for a stochastic POMP model, we can use the sequential Monte Carlo method, also called the particle filter.

We present the main idea behind the calculation of the likelihood via the particle filter below. This is based on the work of Kitagawa [1987], Doucet et al. [2001], Arulampalam et al. [2002], King [2016b]. The first step to the calculation of the likelihood is to factor it in the following manner:

$$
\begin{aligned}
\mathcal{L}(\theta) &= \mathbb{P}(y_{1:n}^* | \theta) \\
&= \mathbb{P}(Y_n = y_n^* | Y_{1:n-1} = y_{1:n-1}^*, \theta) \mathbb{P}(Y_{1:n-1} = y_{1:n-1}^*, \theta) \\
&= \mathbb{P}(Y_n = y_n^* | Y_{1:n-1} = y_{1:n-1}^*, \theta) \mathbb{P}(Y_{n-1} = y_{n-1}^* | Y_{1:n-2} = y_{1:n-2}^*, \theta) \mathbb{P}(Y_{1:n-2} = y_{1:n-2}^*, \theta)
\end{aligned}
$$

Continuing on in this manner yields,

$$
\begin{aligned}
\mathcal{L}(\theta) &= \prod_{i=1}^{n} \mathbb{P}(Y_i = y_i^* | Y_{1:i-1} = y_{1:i-1}^*, \theta) \\
&= \prod_{i=1}^{n} \sum_{x_i} \mathbb{P}(Y_i = y_i^* | X_i = x_i, \theta) \mathbb{P}(X_i = x_i | Y_{1:i-1} = y_{1:i-1}^*, \theta)
\end{aligned}
$$

We call the distribution $\mathbb{P}(X_i = x_i | Y_{1:i-1} = y_{1:i-1}^*, \theta)$ the prediction distribution since this provides the distribution of the state X_i given $y_{1:i-1}^*$, all the past data. We also introduce the filter distribution, $\mathbb{P}(X_i = x_i | Y_{1:i} = y_{1:i}^*, \theta)$ which is the distribution of the state given all the past data and the current data. The prediction distribution

can be written in terms of the filter distribution and the process model distribution,

$$\mathbb{P}(X_i = x_i | y_{1:i-1}, \theta) = \sum_{x_i = x_{i-1}} \mathbb{P}(X_i = x_i | X_{i-1} = x_{i-1}, \theta) \mathbb{P}(X_{i-1} = x_{i-1} | Y_{1:i-1} = y^*_{1:i-1}, \theta)$$

We can also write the filter distribution in terms of the prediction distribution,

$$
\begin{aligned}
\mathbb{P}(X_i = x_i | Y_{1:i} = y^*_{1:i}, \theta) &= \mathbb{P}(x_i | y_i, y_{1:i-1}, \theta) \\
&= \frac{\mathbb{P}(y_i | x_i, \theta) \mathbb{P}(x_i | y_{1:i-1}, \theta)}{\sum_{x_i} \mathbb{P}(y_i | x_i, \theta) \mathbb{P}(x_i | y_{1:i-1}, \theta)}
\end{aligned}
$$

Thus at time t_{i-1}, given a set of J points from the filtering distribution $\left\{ x^F_{i-1,j} \right\}_{j=1}^{J}$, we can simulate from this set to the next time point t_i to obtain a set of J values $\left\{ x^P_{i,j} \right\}_{j=1}^{J}$ which would come from the prediction distribution,

$$x^P_{i,j} \sim \mathbb{P}(X_i = x | X_{i-1} = x^F_{i-1,j}, \theta), \qquad j = 1, \ldots, J.$$

A set of points from the filtering distribution can then be obtained by re-sampling from $\left\{ x^P_{i,j} \right\}_{j=1}^{J}$ with weights proportional to $\mathbb{P}(Y_i = y^*_i | X_i = x^P_{i,j}, \theta)$. Thus we can derive an estimate of the conditional likelihood at the ith time index of observation,

$$\mathcal{L}_i(\theta) \approx \frac{1}{J} \sum_j \mathbb{P}(Y_i = y^*_i | X_i = x^P_{i,j}, \theta).$$

Sequentially solving for the conditional likelihoods in this manner until the last time index yields an unbiased estimate of likelihood (Bretó et al. [2009], King [2016b,a]),

$$\mathcal{L}(\theta) = \prod_{i=1}^{n} \mathcal{L}_i(\theta) \approx \prod_{i=1}^{n} \left[\frac{1}{J} \sum_j \mathbb{P}(Y_i = y^*_i | X_i = x^P_{i,j}, \theta) \right].$$

The particle filter yields stochastic estimates of the likelihood which we would like to maximize when we fit the model to data. There are different methods to

maximize stochastic likelihood estimates to get the MLE. In this project we used the maximization via iterated filtering algorithm of Ionides et al. [2015, 2006], Ionides [2011] and implemented in the R package pomp (King et al. [2017], Ionides et al. [2016]). The mechanics of this algorithm is beyond the scope of this thesis but is discussed in Ionides et al. [2016], King [2016b].

2.7 Akaike information criterion (AIC)

The Akaike information criterion (AIC) (Akaike [1974]) is a criterion from information theory that facilitates the selection of a preferred model among a suite of different models. The AIC measures the relative "goodness of fit" among the competing models but also includes a penalty for overfitting. As discussed earlier, the likelihood gives a sense of how probable it is that the data is taken from a given model (in our case the entire POMP model including both the process and observation model) with parameter values given by the parameter vector θ. The maximized log-likelihood is a measure of the bias in the model, but it does not take into account the variability in the model. Variability and uncertainty increases in the model with the number of free model parameters. All things being equal, the more free parameters a model has the more variability it will have. There is usually a trade off between increasing the number of free parameters and variability (including error introduced by these free parameters) in a model. The AIC selects the optimum among competing models by balancing both the bias and the variability. For the maximum likelihood case, it is defined by the following equation,

$$\text{AIC} = -2(\ell(\theta^*)) + 2K,$$

where $\ell(\theta^*)$ is the maximized log-likelihood of the model and K is the number of free parameters used to fit the model.

To compare models using the AIC, the same data should be used to fit the models, since comparison cannot be made otherwise. The most preferred among the competing models corresponds to that with the least AIC score. The value of the AIC scores by themselves does not mean anything and this criterion should only be used to help with ranking models (Akaike [1974, 1981]).

2.8 Summary

In this chapter, we presented the mathematical preliminaries for this study, including the basic properties of the simple deterministic SEIR model (well-posedness, invariance in the simplex, existence and stability of the disease-free equilibrium, derivation of the basic reproduction number, existenceand stability of the endemic equilibrium). We also presented the steps to build this model up into a stochastic SEIR model with a school-term forcing function using counting processes. This will be the basis of many models discussed in the next chapter.

In this chapter we also discussed the likelihood function, calculation of the likelihood using the particle filter, the maximum likelihood estimates of parameters of a model and the AIC. In the next chapter we will use these concepts to fit some of the focus measles models. Later on we will also use the AIC for model comparison.

3

MEASLES MODELS

In this chapter, we review some of the well-known models of measles transmission and present a detailed discussion of the four models that we focus on in this thesis: (1) the deterministic SEIR model with school-term forcing, (2) the stochastic SEIR model with school-term forcing, (3) the Bjornstad et al. [2002] TSIR model, and (4) the He et al. [2010] model which includes school-term forcing, the cohort effect and infections from outside the population.

We begin by reviewing the literature on the history of measles transmission models in section 3.1. In section 3.2, we describe the open questions that have motivated this review of measles models, and describe the work that has been done for this project. In section 3.3, we present a description of the London data set that we used for comparing the different models. In sections 3.4–3.7, we discuss each of the four focus models. For each model, we present the assumptions inherent in each model, the method by which the model is fitted to the data, the maximum likelihood estimates of the model parameters, and its strengths and weaknesses. A summary of the different properties of the four focus models is presented in Table 3.5.

3.1 Literature review

Measles is the poster child for disease modeling. The disease is very infectious and most people who recover from it gain lifelong immunity. Thus in the pre-vaccine era, most adults would have already had measles and were effectively removed from the transmission pool. Among children it is reasonable to assume homogeneous contact rates that varies with the opening and closing of schools. The standard SEIR model framework with the addition of school-term forcing is appropriate for modeling this.

Measles transmission dynamics typically exhibits regular periodic behaviors in the pre-vaccine era and there is a rich collection of data on the disease. In this study, we focus on pre-vaccine era data in London, UK where the disease exhibited annual dynamics before 1949 which gave way to more biennial dynamics from 1949 until the start of the vaccine era in 1968. It has been shown that the simple SEIR formulation with the appropriate seasonal forcing can capture (at least qualitatively) these annual and biennial dynamics of measles (Earn et al. [2000a]).

Many models including a range of empirical and mechanistic models have been formulated to explain the physical and biological aspects of measles transmission (Schwartz [1985], Olsen and Schaffer [1990], Sugihara et al. [1990], Rand and Wilson [1991], Kendall et al. [1994], Ellner et al. [1995, 1998], Grenfell and Harwood [1997], Earn et al. [2000b], Mollison and Ud Din [1993], He et al. [2010], Cauchemez et al. [2008], Grenfell et al. [2002]). Below we provide a short discussion of some of the well-known mathematical studies of measles transmission.

Bartlett [1957] and Bartlett [1960]

In a study by Bartlett [1957] of measles in pre-vaccination era UK, it was noted that there were two main types of dynamics that could be observed: periodic endemic dynamics in large populations, and episodic outbreaks in smaller communities. Bartlett [1957] noted that the latter is due to random extinctions or "fadeouts" of the disease

which are more likely to occur in smaller communities. He derived a threshold population size (also known as the critical community size) that differentiates this large and small population dynamics, estimating that it was about 200,000–250,000. In a later study, Bartlett [1960] estimated that this threshold is between 250,000–300,000 using data from the the United States. Existing estimates of the critical community size for measles now range from 250,000-500,000 (Bjornstad et al. [2002],Keeling and Grenfell [1997]).

London and Yorke [1973]

In Chapter 2, we noted that the standard deterministic SEIR model with no seasonal forcing cannot yield the sustained periodic outbreaks of measles expected in the pre-vaccine era. This was pointed out as a problem by London and Yorke [1973] for modeling measles childhood diseases. The authors proposed that employing variable contact rates could account for the observed seasonal variation in number of cases, and that the variable contact rates could arise from the closer contacts among children during the school terms. The authors presented an SEIR-type differential equation model with variable monthly contact rates and delays due to the disease incubation and infectious periods. The models were fitted to childhood disease data in the US and their measles model was able to reproduce the biennial dynamics of pre-vaccine measles in the New York City dataset they used (also observed in the later part of the London dataset used in this thesis.) The contact rates calculated for measles, chickenpox and mumps all reflected higher contact rates during the school-term.

The authors also studied the effect of the latent period and infectious period on the dynamics of the disease. They observed that a shorter infectious period led to more disease extinctions. Additionally, increasing the length of the infectious period could change the model dynamics from exhibiting a biennial cycle to an annual cycle.

Schenzle [1984]

Schenzle [1984] proposed a model that accounted for the age-structure of a population to reflect the variation in contact rates between different age classes. His model is called the realistic age-structured model (RAS model). It was derived as an SEIR system consisting of PDEs that was simplified into an SEIR system of ODEs with multiple age classes. From this an approximate stochastic formulation of the model was derived. The RAS model incorporates a changing contact rate among school-aged children due to schools opening and closing. Like the London and Yorke [1973] model, this model successfully captured the biennial epidemics of measles in England and Wales. Age-specific prevalence curves derived from this model were also closer to the data than those derived from mass action models.

Keeling and Grenfell [1997]

The stochastic dynamics of the RAS model were found to be inadequate for modeling transmission in populations with less than a million people (Keeling and Grenfell [1997]). It generated more fade outs than was observed in recorded data. To rectify this, Keeling and Grenfell [1997] proposed a pulsed realistic age-structured model (PRAS model). In this model the standard assumption of constant transition rates (and therefore exponentially distributed times) were replaced with a normally distributed incubation and infection period (Keeling and Grenfell [1997]). This change resulted in more intense "pulses" of infection, hence the name of the model. The PRAS model yielded a better fit in estimating the critical community size.

Finkenstadt et al. [1998]

The effect of covariates such as birth rates and the population sizes on the dynamics of measles were explored by Finkenstadt et al. [1998]. Using the PRAS model of Keeling and Grenfell [1997], the authors found that small population sizes corresponded to

irregular disease patterns while larger population sizes corresponded to more regular and predictable endemic patterns of disease. The authors also showed that birth rates affected the cyclical dynamics of the disease. In particular lower birth rates led to a transition from regular biennial epidemics to more annual epidemics in the model simulations corresponding to what has been observed in real data.

Bjornstad et al. [2002]

The time-series SIR (TSIR) model constructed by Bjornstad et al. [2002] is a well-known measles model as it is one of the first to explore not only the endemic dynamics of measles in the major cities but also the episodic dynamics in small cities, where there are epidemics followed by disease extinction and reintroduction (Cliff et al. [1993]). It does this by first reconstructing the dynamics of the number of susceptible from recorded birth and incidence data using the Susceptible Reconstruction Algorithm (Finkenstädt et al. [2000]). It does not take into account age classes but allows for a seasonal forcing function and non-linearity in contact rates. The TSIR model is a discrete time stochastic SEIR model with a time step that has to match up with the time between data points. This allowed fitting the model to data. Further details on this model will be discussed later on in this chapter.

Cauchemez and Ferguson [2007]

Cauchemez and Ferguson [2007] proposed a continuous-time, discrete-state Markov chain SIR model for measles. Unlike the Bjornstad et al. [2002] TSIR model, this model allowed for the data and generation time of the disease to be at different time scales and for data to be collected at irregular intervals. An approximation of the model was derived using a Cox-Ingersoll-Ross (CIR) diffusion process and Metropolis-Hastings Markov chain Monte Carlo (MCMC) was used to estimate parameters using likelihood. This method was found to be successful in simulation

studies and the authors concluded that generation time could be well estimated if the observation interval is less than 2.5 times the generation time. However the authors found that applying the method to fit measles pre-vaccine era data was unsuccessful at estimating the known generation times of measles. They suggested that simple models for measles with homogeneous contact rates may be inadequate.

He et al. [2010]

The He et al. [2010] model is another model that we explore in further detail later in this thesis. This model is a continuous time Markov Chain SEIR model. It uses actual data on the demography of populations as covariate, and allows for extra demographic stochasticity and imported infections. Both school-term forcing and the cohort effect are included in the model. This model was formulated and fitted to UK data using new plug-and-play (i.e. simulation-based) inference techniques called maximization via iterated filtering, which is implemented in R software package *pomp* (King et al. [2016, 2017]). This study yielded a model for the London data (and other UK cities) with very good likelihood and AIC relative to other modeling studies. However the authors also noted that they obtained values for the basic reproduction number that are higher than what would be estimated based on other forms of data (e.g. serological surveys). This again may be due to the use of a model with homogeneous contact rates, and is further motivation for the review of existing measles models in this thesis.

3.2 Summary of the work done for this thesis

In this thesis, we perform a detailed study of four different well-known models for pre-vaccination era measles: the deterministic SEIR model with school-term forcing, the stochastic SEIR model with school-term forcing, the Bjornstad et al. [2002] TSIR

model, and the He et al. [2010] model. This involved studying the papers on modeling measles, presenting the assumptions of the four different models, writing the code to numerically solve or simulate each of these models, applying appropriate statistical inference techniques to find the maximum likelihood estimates of the parameters of each model, and assessing the strengths and weaknesses of each model. We compared these models based on how they fit the pre-vaccine era London data set using the likelihood, Akaike information criterion and residuals for comparison. Finally we also compared how these models perform when extended to the transition period and to the vaccine era. The goal of this project is to review the developments in modeling measles in order to gain a better understanding of the major mechanisms that need to be included in pre-vaccine era measles transmission models. This review should also help with future work on identifying the shortcomings of simple extensions of these models into the vaccine era.

3.3 Description of the dataset

We used data on reports of measles from the pre-vaccine era in London, UK. The data is obtained from url: `http://kingaa.github.io/pomp/vignettes/twentycities.rda`. This weekly data from the start of 1944 to the end of 1963 had been obtained from the UK Registrar General (OPCS) under the national notification program which started in 1944. Aside from the measles reports data, we also used the annual birth rates and annual population data for London from this data set. For this project we aggregated the measles reports biweekly and used a spline of the population and birth data to obtain biweekly estimates of these covariates as shown in figure 2.1.

To compare models of measles in the transient and vaccine era, data on reports of measles in London from 1944 to 1994 (including the vaccine era), and the corresponding data on demographic covariates, was obtained from the International

Infectious Disease Data Archive (IIDDA) at url: `http://lalashan.mcmaster.ca/`
`theobio/IIDDA/index.php/Main_Page`.

3.4 Deterministic SEIR model with school-term forcing

The first focus model of this thesis is the deterministic SEIR model with school-term forcing. Here we take the deterministc model with forcing discussed in Chapter 2 and allow it to have varying birth rates and population sizes so that it can reflect the actual London data on birth rates and population sizes.

3.4.1 Description

In section 2.2 we presented a deterministic SEIR model with a changing transmission rate due to school-term forcing as one way to improve the standard SEIR model to enable it to exhibit the periodic dynamics seen in the data (Earn et al. [2000a]). Here, we again use a periodic forcing function $B(t)$ for the time-dependent transmission rate which reflects the opening and closing of school terms. Additionally we introduce a time-dependent birth rate $\mu_b(t)$. Since we are focusing on homogeneous models, we do not have a separate class for infants and very young children who are expected to have low contact rates. Thus to accommodate this we instead introduce a delay of $\tau = 4$ years to reflect a delay in entry from birth into the susceptible class. We also assume that the vaccine coverage is now a time-dependent function $v(t)$. A fraction $v(t)$ of births at time t are assumed to be vaccinated and go directly to the recovered/removed class. When we consider only the pre-vaccine era we set $v(t) = 0$.

The equations for the model are given as the system (3.1)—(3.4):

$$S' = \mu_b(t - \tau)(1 - v(t - \tau)) - B(t)\frac{SI}{N} - \mu S \tag{3.1}$$

$$E' = B(t)\frac{SI}{N} - (\lambda + \mu)E \tag{3.2}$$

$$I' = \lambda E - (\gamma + \mu)I \tag{3.3}$$

$$R' = \gamma I - \mu R + \mu_b(t - \tau)v(t - \tau) \tag{3.4}$$

We require the following assumptions on the initial conditions of the system,

$$S(0) \geq 0, \quad E(0) \geq 0, \quad I(0) \geq 0, \quad R(0) \geq 0 \tag{3.5}$$

and the non negativity of model parameters. The school-term forcing function $B(t)$ is defined to be,

$$B(t) = \begin{cases} \beta\,(1 + 2(1 - p)a), & \text{during the school term,} \\ \beta\,(1 - 2pa), & \text{during vacation,} \end{cases} \tag{3.6}$$

where β is the average transmission rate in one year, a is the amplitude parameter, and p the proportion of the school days for which the school-term is in session. As before, the days included in the school term for the school-term forcing function is shown in table 2.1.

3.4.2 Fitting the model to data

To fit the model to data, we assume that the number of new cases of measles occurring within the time interval $[t_1, t_2]$ is $C(t_1, t_2) = \int_{t_1}^{t_2} \gamma\, I(t)dt$ which gives the number of transitions from I to R in that time interval. We note that a different way to obtain the number of cases would be to define $C(t_1, t_2) = \int_{t_1}^{t_2} B(t)\frac{S(t)I(t)}{N}dt$ which yields the number of transitions from S to E. We choose to use the latter definition of

$C(t_1, t_2)$ because we assume that reports are likely to occur after reaching the infectious stage, then after reporting to a doctor these individuals are more aware and less likely to contribute to new infections by (voluntarily or involuntarily) quarantining themselves. We also assume that the cases have a reporting probability of $\rho \in (0, 1)$. Following He et al. [2010], we assume that the reports of measles within the time interval $[t_1, t_2]$ have mean $\rho C(t_1, t_2)$, and variance $\rho(1-\rho)C(t_1, t_2) + (\psi \rho C(t_1, t_2))^2$. The first term $\rho(1-\rho)C(t_1, t_2)$ of this variance corresponds to the standard variance of a binomial distribution. The second term $(\psi \rho C(t_1, t_2))^2$ allows for over-dispersion. Thus ψ is an over-dispersion parameter and the reports have an over-dispersed binomial form. Again following He et al. [2010] we set the reports to have the following distribution,

$$\text{reports} \sim \text{Normal}\left(\rho C(t_1, t_2), \rho(1-\rho)C(t_1, t_2) + (\psi \rho C(t_1, t_2))^2\right). \tag{3.7}$$

The model was encoded as a pomp object using the *pomp* package from the Comprehensive R Archive Network (CRAN) (King et al. [2016, 2017]). The pomp object was provided with a vector field type deterministic skeleton given by the differential equations (3.1)—(3.4) and a measurement density and measurement process following (3.7) (using a binned normal distribution). State processes were initiated at 1940 (the start of the year 1940 corresponds to $t = 0$) and fitted to the London data from 1944–1964.

In Table 3.1 we present a a description of the parameters of the model and a list of which parameters were fixed or estimated. We note that here μ now represents the death rate minus the net immigration rate of the population. Additionally, the values of the incubation and recovery rates were fixed so that the latent and infectious periods were fixed at eight and five days respectively. The basic reproduction number was estimated using the mean transmission rate and equation (2.14). Parameter fitting was performed using the trajectory matching algorithm in the *pomp*

package and the subplex algorithm for optimization. A global search for the maximum likelihood estimates of parameters was performed. For each of the estimated parameters, trajectory matching was initiated from different points in its range using the *pomp* function profileDesign. The likelihood was maximized over all other estimated parameters while holding the focal parameter fixed.

Table 3.1: Descriptions and values of parameters for the deterministic SEIR model at the MLE

Symbols	Description	Range	Value	Fixed/Estimated
μ	Death minus net immigration rate	$[0,\infty)$	0.02 yr^{-1}	Estimated
τ	Delay of entry into susceptible class	$[0,\infty)$	4 yr	Fixed
σ	Incubation rate	$[0,\infty)$	45.66 yr^{-1}	Fixed
γ	Recovery rate	$[0,\infty)$	73.05 yr^{-1}	Fixed
R_0	Basic reproduction number	$[0,\infty)$	31.19	Estimated
a	Amplitude of seasonality	$[0,1]$	0.38	Estimated
d	Recruitment delay	$[0,\infty)$	4 yrs	Fixed
ρ	Reporting probability	$[0,1]$	0.46	Estimated
ψ	Reporting overdispersion	$[0,\infty)$	0.90	Estimated
$N(0)$	Initial population size	$[0,\infty)$	2,445,368	Fixed
$S(0)/N(0)$	Initial fraction of susceptibles	$[0,1]$	0.03	Fixed
$E(0)/N(0)$	Initial fraction of exposed	$[0,1]$	5.17×10^{-5}	Fixed
$I(0)/N(0)$	Initial fraction of infected	$[0,1]$	5.14×10^{-5}	Fixed
$R(0)/N(0)$	Initial fraction of recovered	$[0,1]$	0.97	Fixed

3.4.3   Model strengths and weaknesses

At the MLE, this model can exhibit the annual and biennial endemic dynamics seen in the London data. We expect however that this model cannot exhibit the episodic dynamics in smaller cities. Such dynamics require the variability in the process model for small populations, which this model cannot recreate because it is deterministic.

There is a gap between the theoretical model and empirical data, because with this ODE model we assume a continuous state model in which the population can be divided into infinitesimally small fractions. Additionally, similar to the findings of He et al. [2010], this model resulted in relatively high R_0 values than have been found from estimating R_0 using serological data (Edmunds et al. [2000]).

3.4.4 Summary of the model

The deterministic model with school-term forcing has the following assumptions and properties:

- The model is a deterministic system of differential equations with continuous time and continuous state variables given by (3.1)–(3.4).

- Cases can be aggregated over any time period by integrating over the number of transitions from I to R.

- Report are assumed to have an over-dispersed binomial distribution approximated by equation (3.7) with constant reporting probability ρ and over-dispersion parameter ψ.

- Annually changing transmission rates assumed to be due to school-term forcing and given by (3.6).

- Actual London birth rates and population used as covariates.

- Entry into the susceptible class is assumed to have a delay of four years corresponding to the delay in school entry.

- The continuous states do not allow for actual fadeouts of the disease (the $E(t)$ and $I(t)$ states go to very low values but cannot go to zero).

3.5 Stochastic SEIR model with school-term forcing

The deterministic SEIR model from the previous section can be extended into a stochastic model using a similar extension described in section 2.3. A schematic of the stochastic SEIR model is shown in Figure 3.1.
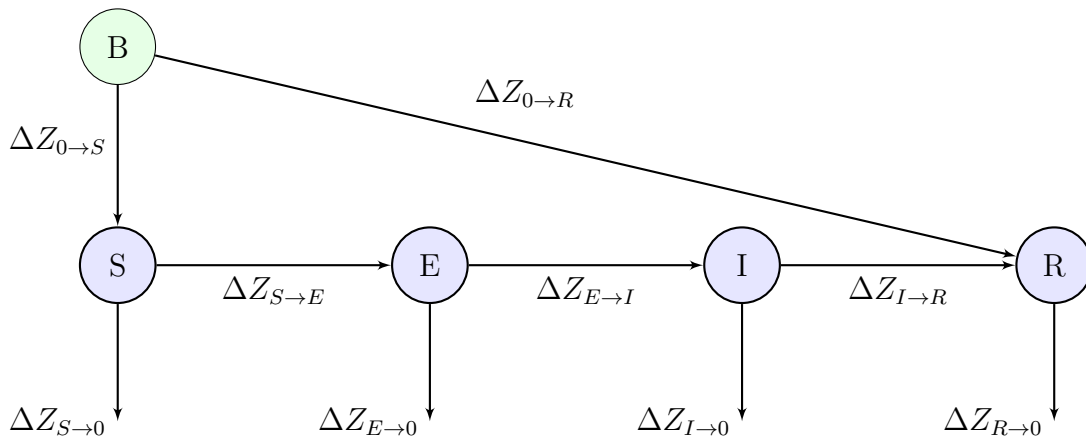
Figure 3.1: Schematic of the stochastic SEIR model with school-term forcing

3.5.1 Description

As in section 2.3, we use the notation $X(t) = (S(t), E(t), I(t), R(t))$ and $Z_{A \to B}(t)$ be the total number of individuals that have transitioned from compartment A to B in the time interval $[0, t]$. Let $t > 0$ and $h > 0$. We define $\Delta Z_{A \to B}(t, t + h) = Z_{A \to B}(t + h) - Z_{A \to B}(t)$ to be the change in the $Z_{A \to B}$ counting process over the time interval $[t, t + h]$. Let $N(t) = S(t) + E(t) + I(t) + R(t)$. We specify the transition

probabilities due to infections and recovery as follows,

$$\mathbb{P}\left(\Delta Z_{S\to E}(t,t+h)=1\mid X(t)\right) \;=\; \frac{B(t)\,S(t)I(t)}{N(t)}\,h+o(h) \qquad (3.8)$$

$$\mathbb{P}\left(\Delta Z_{E\to I}(t,t+h)=1\mid X(t)\right) \;=\; \sigma\,E(t)h+o(h) \qquad (3.9)$$

$$\mathbb{P}\left(\Delta Z_{I\to R}(t,t+h)=1\mid X(t)\right) \;=\; \gamma\,I(t)h+o(h) \qquad (3.10)$$

We use the same $B(t)$ given in (3.6) as used in the deterministic version of this model. Table 2.1 shows the start and end dates of days for which the school-term is in session in London, used in the school-term forcing function.

We use time-varying birth rates $\mu_b(t)$ given in the London data, and a constant death rate of μ. Thus births into the S and R components, and deaths are given by the following equations,

$$\mathbb{P}\left(\Delta Z_{O\to S}(t,t+h)=1\mid X(t)\right) \;=\; \mu_B(t-\tau)\,(1-v(t-\tau))\,h+o(h) \quad (3.11)$$

$$\mathbb{P}\left(\Delta Z_{O\to R}(t,t+h)=1\mid X(t)\right) \;=\; \mu_b(t-\tau)\,v(t-\tau)\,h+o(h) \qquad (3.12)$$

$$\mathbb{P}\left(\Delta Z_{S\to O}(t,t+h)=1\mid X(t)\right) \;=\; \mu\,S(t)h+o(h) \qquad (3.13)$$

$$\mathbb{P}\left(\Delta Z_{E\to O}(t,t+h)=1\mid X(t)\right) \;=\; \mu\,E(t)h+o(h) \qquad (3.14)$$

$$\mathbb{P}\left(\Delta Z_{I\to O}(t,t+h)=1\mid X(t)\right) \;=\; \mu\,I(t)h+o(h) \qquad (3.15)$$

$$\mathbb{P}\left(\Delta Z_{R\to O}(t,t+h)=1\mid X(t)\right) \;=\; \mu\,R(t)h+o(h) \qquad (3.16)$$

We again assume that multiple transitions within the time interval $[t,t+h]$ are $o(h)$. Though there is a delay in model entry, this does not affect the Markov property of the model. This is because the model's birth term is from recorded number of births and this is fixed and independent of model states. Thus this delay in births can be viewed as an independent at any point in time, thus keeping the Markov property intact. The changes in the sizes of the compartments and subcompartments over the

time interval $[t, t + h]$ interval is given by the following equations.

$$\Delta S = \Delta Z_{O \to S}(t, t + h) - \Delta Z_{S \to E}(t, t + h) - \Delta Z_{S \to O}(t, t + h) \qquad (3.17)$$

$$\Delta E = \Delta Z_{S \to E}(t, t + h) - \Delta Z_{E \to I}(t, t + h) - \Delta Z_{E \to O}(t, t + h) \qquad (3.18)$$

$$\Delta I = \Delta Z_{E \to I}(t, t + h) - \Delta Z_{I \to R}(t, t + h) - \Delta Z_{I \to O}(t, t + h) \qquad (3.19)$$

$$\Delta R = \Delta Z_{O \to R}(t, t + h) + \Delta Z_{I \to R}(t, t + h) - \Delta Z_{R \to O}(t, t + h) \qquad (3.20)$$

The stochastic scheme used to simulate this process is the Euler-multinomial scheme which is described in the following references: King [2016b], He et al. [2010], Bretó et al. [2009].

3.5.2 Fitting the model to data

For this model we assume, as in the deterministic case, that cases are only reported during the transition from I to R. Thus for this model we set $C(t_1, t_2) = Z_{I \to R}(t_2) - Z_{I \to R}(t_1)$. Given $C(t_1, t_2)$, the number of actually reported cases within the time period $[t_1, t_2]$ are also assumed to be given by (3.7) (over-dispersed binomial form) with observation probability ρ and over-dispersion parameter ψ.

The model was encoded as a pomp object using the *pomp* package (King et al. [2016, 2017]). The pomp object was provided with a process model given by (3.8)–(3.16) and (3.17)–(3.20), and a measurement model following (3.7). State processes were again initiated at 1940 and fitted to the London data from 1944–1964. Parameter fitting was performed using the maximization via iterated filtering algorithm in the *pomp* package (King et al. [2016, 2017]). A search for the maximum likelihood estimates of parameters was performed by initiating value at the MLEs of the deterministic model.

In Table 3.2 we present a a description of the parameters of the stochastic SEIR model and a list of which parameters were fixed or estimated. Once again the param-

eter μ here represents death rate minus the net immigration rate of the population. Additionally, the values of the incubation and recovery rates were again fixed so that the latent and infectious periods were fixed at eight and five days respectively. The basic reproduction number is also an approximation using the mean transmission rate and (2.14).

Table 3.2: Descriptions and values of parameters for the stochastic SEIR model with school term forcing at the MLE

Symbols	Description	Range	Value	Fixed/Estimated
μ	Death minus net immigration rate	$[0, \infty)$	8.4×10^{-4} yr^{-1}	Estimated
σ	Incubation rate	$[0, \infty)$	45.66 yr^{-1}	Fixed
γ	Recovery rate	$[0, \infty)$	73.05 yr^{-1}	Fixed
R_0	Basic reproduction number	$[0, \infty)$	28.02	Estimated
a	Amplitude of seasonality	$[0, 1]$	0.33	Estimated
τ	Recruitment delay	$[0, \infty)$	4 yrs	Fixed
ρ	Reporting probability	$[0, 1]$	0.40	Estimated
ψ	Reporting over-dispersion	$[0, \infty)$	0.57	Estimated
$N(0)$	Initial population size	$[0, \infty)$	2,445,368	Fixed
$S(0)/N(0)$	Initial fraction of susceptibles	$[0, 1]$	0.03	Fixed
$E(0)/N(0)$	Initial fraction of exposed	$[0, 1]$	5.17×10^{-5}	Fixed
$I(0)/N(0)$	Initial fraction of infected	$[0, 1]$	5.14×10^{-5}	Fixed
$R(0)/N(0)$	Initial fraction of recovered	$[0, 1]$	0.97	Fixed

3.5.3 Model strengths and weaknesses

At the MLE, this model can also exhibit both the annual and biennial endemic dynamics seen in the London data. This model is expected to also do better than the deterministic model at exhibiting the dynamics in smaller cities. Unlike the ODE system describing the previous deterministic SEIR model, this model uses discrete

states which allow for fadeouts of the disease. However the model does not include infections from outside the population and thus fadeouts of the disease in this model would correspond to permanent extinction, which is not what we expect in small cities. Additionally, this model is more computationally expensive to simulate and fit than the deterministic model. This model also exhibits the relatively higher R_0 values than would be expected from serological studies of measles.

3.5.4 Summary of the model

The stochastic SEIR model with school-term forcing involves the following assumptions and has the following properties:

- The model is a continuous time Markov chain model with discrete states and transitions given by (3.8)–(3.16) and (3.17)–(3.20).

- Cases can be aggregated over any time period by adding the number of transitions from I to R.

- Report are assumed to have an over-dispersed binomial distribution given by equation (3.7) with constant reporting probability ρ and over-dispersion parameter ψ.

- Annually changing transmission rates assumed to be due to school-term forcing and given by (3.6).

- Actual London birth rates and population were used as covariates.

- Entry into the susceptible class is assumed to have a delay of four years corresponding to the delay in school entry.

- The model includes stochasticity in the transitions which allows for more variability in dynamics.

- Discrete states allow this model to display stochastic fadeout in small populations, but a fadeout would correspond to permanent extinction of the disease.

3.6   Bjornstad et al. [2002] model

The Bjornstad et al. [2002] TSIR model takes a different approach to modeling the dynamics of measles. This discrete time, process involves first reconstructing the dynamics of the unobserved susceptible class, using cumulative cases of measles and cumulative number of births. This algorithm is described below.

3.6.1   Susceptible reconstruction algorithm (SRA)

In the susceptible reconstruction algorithm (SRA) we assume that we can model the disease using discrete time steps of length equal to the generation time of the disease. Thus in the case of measles, we assume that each time step is two weeks long. At each step all infections are assumed to be new and all previous infections are assumed to have recovered (thus making the incidence of the disease equal to the prevalence). By also assuming that during the pre-vaccine era, all hosts would have been infected by measles at some point, we can derive the following relationship between S_t, the number of susceptibles at time t, and I_t, the number of infected people at time t (Finkenstädt et al. [2000]):

$$S_t = B_{t-d} + S_{t-1} - I_t + u_t,$$

Here B_t is the number of susceptible births at time t (if there is vaccination then it corresponds to the unvaccinated newborns). The parameter d corresponds to a delay that allows maternal immunity to wane, which for measles is approximately 16 weeks (Finkenstädt et al. [2000], Anderson and May [1991]). Since we assume that we have a time step of two weeks, we set $d = 8$. Additionally, u_t be an additive error

factor, with $E(u_t) = 0$ and $\text{Var}(u_t) = \sigma_u^2$.

We also assume that the actual number of measles cases at any time is proportional to the number of reported cases, i.e. $I_t = \rho_t C_t$ where C_t is the number of reported cases and $\rho_t \geq 1$ is a random variable such that $E(\rho_t) = \rho$. We note that for this model, ρ is the inverse of how this has been used in the other models where the parameter ρ was the actual reporting probability.

The equation for the susceptibles can be rewritten as (Finkenstädt et al. [2000]):

$$S_t = B_{t-d} + S_{t-1} - \rho_t C_t + u_t \tag{3.21}$$

We now introduce a new variable $Z_t = S_t - \bar{S}$ where $\bar{S} = E(S_t)$ so that $E(Z_t) = 0$. Clearly S_t and Z_t have similar recursive relations. By subtracting $\bar{S}$ from both sides of (3.21) we derive,

$$Z_t = B_{t-d} + Z_{t-1} - \rho_t C_t + u_t. \tag{3.22}$$

Successively iterating the above equation leads to:

$$Z_t = Z_0 + \sum_{i=1}^{t} B_{i-d} - \sum_{i=1}^{t} \rho_i C_i + \sum_{i=1}^{t} u_i \tag{3.23}$$

This last equation is free from the initial autoregressive form of S_t and Z_t in (3.21) and (3.23). We now define the following variables,

$$X_t = \sum_{i=1}^{t} C_i \tag{3.24}$$

$$Y_t = \sum_{i=1}^{t} B_{i-d} \tag{3.25}$$

$$U_t = \sum_{i=1}^{t} u_i \tag{3.26}$$

$$R_t = \sum_{i=1}^{t} (\rho_i - \rho) C_i \tag{3.27}$$
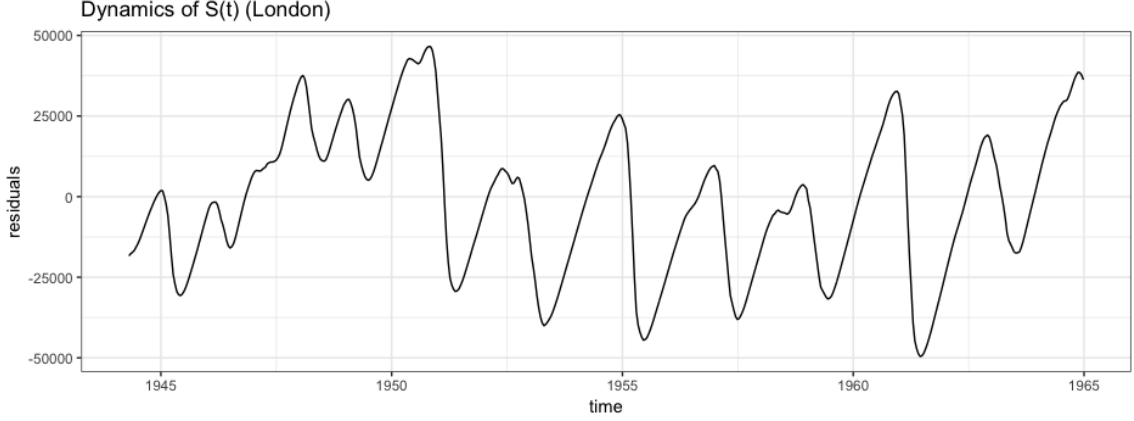
Figure 3.2: Estimates of the variation from equilibrium of the size of the susceptible population to measles (Z_t from (3.23)) for London from 1944–1964 using the SRA.

for $t = 1, ..., n$. Then the following are random-walk processes,

$$R_t = R_{t-1} + (\rho_t - \rho)C_t,$$

$$U_t = U_{t-1} + u_t.$$

With the new notation we have introduced, the regression equation becomes

$$Y_t = -Z_0 + R_t + \rho X_t + Z_t - U_t.$$

Recalling that $E(Z_t) = 0$, under the assumption of a constant reporting rate ($R_t \approx 0$) and noise free data ($U_t \approx 0$) we have a simple linear relation between cumulative births (Y_t) and cumulative incidence (X_t). A point estimate of the average reporting rate is the reciprocal of the estimate of the slope (ρ). Additionally, the regression equation allows us to estimate the Z_t by looking at the residuals to the linear relationship. Figure 3.2 shows the dynamics of Z_t. Since $Z_t = S_t = \bar{S}$, this also yields the dyanmics of the susceptible class S_t about $\bar{S}$.

The assumption of a constant reporting rate can be refined to a constant reporting within each observation interval. In this case, Local Regression Analysis (LRA) can be used to obtain a better fit. The local linear regression is done with a Gaussian

kernel. If the bandwidth of the LRA is too small the resulting fitted values will be the same as the dependent variable. Also, a bandwidth large enough will result in a local regression with the same results as the global linear regression (with the assumption that errors are normally distributed about 0 with constant variance). Thus a bandwidth should be chosen that optimizes the explanatory power of LRA while preserving the cyclicality in the remainder (Finkenstädt et al. [2000]). This is chosen at the intersection of the sum of squared error (SSE) of the local linear regression and the SSE between the global and local linear regression's fitted values.

3.6.2 Description

We define θ_t to be the number of migrant infected hosts at time t and assume that this is Poisson distributed with fixed rate m. We also define a time varying transmission rate such that,

$$\beta_t = \beta_{t+26}. \tag{3.28}$$

Since each time step is two weeks apart, this means that the transmission rate has a one-year period. For the TSIR model, we allow for 26 different values of the transmission rate. This is in contrast to the two different values allowed using the school-term forcing function in the previous two models.

The TSIR model is fully specified by initial conditions I_0 and S_0, and the following system of equations:

$$\lambda_{t+1} = \beta_t (I_t + \theta_t)^\alpha S_t \tag{3.29}$$

$$I_{t+1} \sim \text{Negative Binomial}(\lambda_{t+1}, I_t) \tag{3.30}$$

$$\theta_t \sim \text{Poisson}(m). \tag{3.31}$$

$$S_t = B_{t-d} + S_{t-1} - I_t + u_t \tag{3.32}$$

Here the parameter α is a "mixing parameter" which allows for non-linearity in contact patterns, and u_t is again the additive error term with $E(u_t) = 0$ and $\text{Var}(u_t) = \sigma_u^2$.

3.6.3 Fitting the model to data

Recall that $I_t = \rho_t C_t$ is defined to be the true number of cases at time t and C_t is the number of reported cases. Using the Z_t and ρ_t from the SRA, the model can be fitted to data on measles reports using the reconstructed model below:

$$\lambda_{t+1} = \beta_t(I_t + \theta_t)^\alpha (\bar{S} + Z_t) \tag{3.33}$$

$$I_{t+1} \approx \text{Negative Binomial}(\lambda_{t+1}, I_t) \tag{3.34}$$

$$\theta_t \approx \text{Poisson}(m) \tag{3.35}$$

Given $C(t_1, t_2)$, the number of actually reported cases within the time period $[t_1, t_2]$ are also assumed to be given by (3.7) which has mean and variance in the form of an over-dispersed binomial with observation probability $1/\rho_t$ and over-dispersion parameter ψ. We note that in Bjornstad et al. [2002], the reporting process is not fully specified. In a later paper by Morton and Finkenstädt [2005], the TSIR model is modified and the reporting process is fully specified to attain a better fit, however, that is beyond the scope of this thesis. Additionally, the reporting process we have here is only used in calculation of the model's likelihood and AIC to facilitate comparison with the other model. However this model was not fitted using the given reporting process.

$$\text{reports} \sim \text{Normal}\left(\frac{1}{\rho_t}C(t_1, t_2), \frac{1}{\rho_t}\left(1 - \frac{1}{\rho_t}\right)C(t_1, t_2) + \left(\psi\frac{1}{\rho_t}C(t_1, t_2)\right)^2\right) \tag{3.36}$$

The parameter estimation was done using a modified version of the runTSIR function of the *tsiR* package of CRAN (Becker and Morris [2017]). This fits model to the London data and parameters are estimated in hierarchy as described in Bjornstad et al. [2002] using simple regression (not maximum likelihood using the given reporting model). In order to compute the likelihood and AIC of this model for comparison with the other models, this model was also encoded using the *pomp* package of the Comprehensive R Archive Network (CRAN) (King et al. [2017]). The pomp object had a process model had a discrete time simulator given by (3.29)–(3.32).

Table 3.3 presents a description of the estimated parameters of the Bjornstad et al. [2002] TSIR model.

3.6.4 Model strengths and weaknesses

Being a discrete state stochastic model, the TSIR can exhibit the endemic dynamics of big cities (annual and biennial dynamics) and the episodic dynamics in smaller cities. This model is also more efficient than a continuous time Markov chain model to simulate, and thus more efficient to fit to data.

By using time steps equal to the generation time of the disease, the TSIR model was one of the first models to bridge the gap between theoretical model and empirical data. However, the assumption that the generation interval should be a multiple of the observation interval, and that all infections recover within one time step is a strong assumption. This also constrains the applicability of the model to other disease where the generation time is not in the same scale as the observation periods. The reliance of the model on the SRA may also affect results since this relies on the assumption that everyone eventually gets measles, and since this algorithm is also less reliable in the early part of the time series as the initial conditions are difficult to estimate without allowances at the beginning of the time series (Bjornstad et al. [2002]).

Table 3.3: Descriptions and values of parameters for the Bjornstad et al. [2002] model estimated using the *tsiR* package of CRAN (Becker and Morris [2017])

Symbol	Description	Range	Value	Fixed/Estimated
α	Mixing exponent	$[0, \infty)$	0.978	Estimated
m	Mean number of visiting infectives	$[0, \infty)$	5	Fixed
$\bar{S}$	Mean number of suscepti-bles	$[0, \infty)$	$1.09258.7 \times 10^5$	Estimated
$1/\rho_t$	reporting rate	$(0, 1)$	0.45 (mean)	Reconstructed
ψ	Over-dispersion parameter	$[0, \infty)$	2.55	Estimated
β_1	Transmission rate 1	$[0, \infty)$	1.164170×10^{-5}	Estimated
β_2	Transmission rate 2	$[0, \infty)$	1.230797×10^{-5}	Estimated
β_3	Transmission rate 3	$[0, \infty)$	1.413762×10^{-5}	Estimated
β_4	Transmission rate 4	$[0, \infty)$	1.242972×10^{-5}	Estimated
β_5	Transmission rate 5	$[0, \infty)$	1.210762×10^{-5}	Estimated
β_6	Transmission rate 6	$[0, \infty)$	1.163952×10^{-5}	Estimated
β_7	Transmission rate 7	$[0, \infty)$	1.050347×10^{-5}	Estimated
β_8	Transmission rate 8	$[0, \infty)$	1.237272×10^{-5}	Estimated
β_9	Transmission rate 9	$[0, \infty)$	9.785120×10^{-6}	Estimated
β_{10}	Transmission rate 10	$[0, \infty)$	1.105360×10^{-5}	Estimated
β_{11}	Transmission rate 11	$[0, \infty)$	1.073516×10^{-5}	Estimated
β_{12}	Transmission rate 12	$[0, \infty)$	1.093355×10^{-5}	Estimated
β_{13}	Transmission rate 13	$[0, \infty)$	1.055878×10^{-5}	Estimated
β_{14}	Transmission rate 14	$[0, \infty)$	9.756590×10^{-6}	Estimated
β_{15}	Transmission rate 15	$[0, \infty)$	9.810637×10^{-6}	Estimated
β_{16}	Transmission rate 16	$[0, \infty)$	9.082545×10^{-6}	Estimated
β_{17}	Transmission rate 17	$[0, \infty)$	6.826086×10^{-6}	Estimated
β_{18}	Transmission rate 18	$[0, \infty)$	7.639430×10^{-6}	Estimated
β_{19}	Transmission rate 19	$[0, \infty)$	7.248873×10^{-6}	Estimated
β_{20}	Transmission rate 20	$[0, \infty)$	1.104402×10^{-5}	Estimated
β_{21}	Transmission rate 21	$[0, \infty)$	1.342619×10^{-5}	Estimated
β_{22}	Transmission rate 22	$[0, \infty)$	1.170455×10^{-5}	Estimated
β_{23}	Transmission rate 23	$[0, \infty)$	1.231981×10^{-5}	Estimated
β_{24}	Transmission rate 24	$[0, \infty)$	9.968867×10^{-6}	Estimated
β_{25}	Transmission rate 25	$[0, \infty)$	1.143540×10^{-5}	Estimated
β_{26}	Transmission rate 26	$[0, \infty)$	1.160626×10^{-5}	Estimated

3.6.5 Summary of the model

The Bjornstad et al. [2002] TSIR model involves the following assumptions and has the following properties:

- The model is a discrete time Markov chain model with discrete states and transitions given by (3.29)–(3.32).

- The interval between each discrete time step needs to be equal to the generation time of the disease (two weeks for measles). The latent period is included in this generation time.

- At every time step, all cases are assumed to be new and all cases from the previous time step are assumed to have recovered.

- Report are assumed to have an over-dispersed binomial distribution given by equation (3.36) with constant reporting probability $\frac{1}{\rho_t}$ and dispersion parameter ψ.

- Annually changing transmission rates are assumed with 26 possible values for the transmission rate throughout the year.

- Actual London birth rates and population used as covariates.

- Entry into the susceptible class are delayed by 16 weeks to correspond to the mean time until maternal immunity has been lost.

- The model includes stochasticity in the transitions which allows for more variability in dynamics.

- Discrete states allow this model to display stochastic fadeouts in small populations.

- If $m > 0$, this allows for importation of infections from outside the region.

- The SRA is used to estimate the size of the susceptible pool in the population and a time-varying reporting rate.

- The model allows for nonlinearity in contact patterns.

3.7 He et al. [2010] Model

He et al. [2010] presented a measles model for pre-vaccine era UK cities to demonstrate the plug-and-play methods for fitting partially observed Markov process models to data using maximization via iterated filtering in the R package *pomp* King et al. [2017]. The model is a refinement of the stochastic SEIR model in section 3.5. It includes new features that appear to be important for measles transmission dynamics such as the cohort effect, demographic, extra demographic noise and imported infections (He et al. [2010]).

3.7.1 Description

We again let $Z_{A \to B}(t)$ again be the total number of individuals that have transitioned from compartment A to B in the time interval $[0, t]$. In the He et al. [2010] model, the transition probabilities as follows

$$\mathbb{P}\left(\Delta Z_{S \to E}(t, t+h) = 1 \mid X(t)\right) = \frac{B(t)\, S(t)}{N(t)} \left(I(t) + \iota\right)^\alpha \zeta(t)\, h + o(h) \quad (3.37)$$

$$\mathbb{P}\left(\Delta Z_{E \to I}(t, t+h) = 1 \mid X(t)\right) = \sigma\, E(t)h + o(h) \quad (3.38)$$

$$\mathbb{P}\left(\Delta Z_{I \to R}(t, t+h) = 1 \mid X(t)\right) = \gamma\, I(t)h + o(h) \quad (3.39)$$

Here the time-dependent transmission rate $B(t)$ is again the school-term forcing function from (3.6). The parameter ι is a constant that allows for imported infections from outside the population $N(t)$. The parameter α is a mixing parameter that allows for variation from the standard mass action contact assumption. The function $\zeta(t)$

is a source of multiplicative Gamma white noise in the transmission rate and we assume that it has intensity given by parameter σ_{SE}.

We also define a parameter c to reflect the strength of the cohort effect, which accounts for cohorts of individuals entering the susceptible pool at the same time, possibly due to this cohort all entering school for the first time at a certain time of year. This effect is incorporated into the model by using $\mu_b(t)$, a time-varying birth rate in to the susceptible class, and $\mu_c(t)$, a time-varying birth rate into the removed class. The values of these functions are derived from the actual birth rates $b(t)$ and the vaccine uptake $v(t)$ using the following equations,

$$
\begin{aligned}
\mu_b(t) &= (1-c)\, b(t-\tau)(1-v(t-\tau)) \\
&\quad + c\,\delta(t-t_0) \int_{t-1}^{t} b(t-\tau-s)(1-v(t-\tau-s))\,ds, \quad (3.40) \\
\mu_c(t) &= (1-c)\, b(t-\tau)v(t-\tau) \\
&\quad + c\,\delta(t-t_0) \int_{t-1}^{t} b(t-\tau-s)v(t-\tau-s)\,ds. \quad (3.41)
\end{aligned}
$$

In this equation, τ is again a delay reflecting the amount of time from birth to entry to the susceptible class (usually the time from birth to school-entry.) In He et al. [2010], the vaccine uptake rate $v(t)$ was assumed to be fixed at zero so they only had the $\mu_b(t)$ function in their equations. Here we include $\mu_c(t)$ to allow for the eventual extension of the model to the vaccine era where $v(t)$ is nonzero. Using (3.40), the

births into the S and R components, and deaths are given by the following equations,

$$\mathbb{P}\left(\Delta Z_{O \to S}(t, t+h) = 1 \mid X(t)\right) \;=\; \mu_b(t)h + o(h) \qquad (3.42)$$

$$\mathbb{P}\left(\Delta Z_{O \to R}(t, t+h) = 1 \mid X(t)\right) \;=\; \mu_c(t)h + o(h) \qquad (3.43)$$

$$\mathbb{P}\left(\Delta Z_{S \to O}(t, t+h) = 1 \mid X(t)\right) \;=\; \mu\, S(t)h + o(h) \qquad (3.44)$$

$$\mathbb{P}\left(\Delta Z_{E \to O}(t, t+h) = 1 \mid X(t)\right) \;=\; \mu\, E(t)h + o(h) \qquad (3.45)$$

$$\mathbb{P}\left(\Delta Z_{I \to O}(t, t+h) = 1 \mid X(t)\right) \;=\; \mu\, I(t)h + o(h) \qquad (3.46)$$

$$\mathbb{P}\left(\Delta Z_{R \to O}(t, t+h) = 1 \mid X(t)\right) \;=\; \mu\, R(t)h + o(h) \qquad (3.47)$$

As in the simpler stochastic SEIR model discussed before, we assume that multiple transitions within the time interval $[t, t+h]$ are $o(h)$. The changes in the sizes of the compartments and subcompartments over the time interval $[t, t+h)$ interval is given by the following equations.

$$\Delta S \;=\; \Delta Z_{O \to S}(t, t+h) - \Delta Z_{S \to E}(t, t+h) - \Delta Z_{S \to O}(t, t+h) \qquad (3.48)$$

$$\Delta E \;=\; \Delta Z_{S \to E}(t, t+h) - \Delta Z_{E \to I}(t, t+h) - \Delta Z_{E \to O}(t, t+h) \qquad (3.49)$$

$$\Delta I \;=\; \Delta Z_{E \to I}(t, t+h) - \Delta Z_{I \to R}(t, t+h) - \Delta Z_{I \to O}(t, t+h) \qquad (3.50)$$

$$\Delta R \;=\; \Delta Z_{O \to R}(t, t+h) + \Delta Z_{I \to R}(t, t+h) - \Delta Z_{R \to O}(t, t+h) \qquad (3.51)$$

3.7.2 Fitting the model to data

To fit this model we assume that reporting is done as in the simple stochastic SEIR model with school-term forcing. This means we again assume that cases can only be reported during the transition from I to R. The total number of new cases of measles that can be reported within a given time interval $[t_1, t_2]$ is $C(t_1, t_2) = Z_{I \to R}(t_2) - Z_{I \to R}(t_1)$. Given $C(t_1, t_2)$, the number of actually reported cases within the time period $[t_1, t_2]$ are also assumed to be given by (3.7) which has mean and

variance in the form of an overdispersed binomial with observation probability ρ and overdispersion parameter ψ.

Like the previous models, the model was encoded as a pomp object using the *pomp* package (King et al. [2016, 2017]). The pomp object was provided with a process model given by (3.37)–(3.47) and (3.48)–(3.51), and a measurement model following (3.7). State processes were also initiated at 1940 and fitted to the London data from 1944–1964. Parameter fitting was performed using the maximization via iterated filtering algorithm in the *pomp* package (King et al. [2016, 2017]). A search for the MLE parameters was performed by initiating value at the MLEs given by He et al. [2010] in their paper for the best fit to the London data.

Table 3.4 provides a description of the parameters we obtained from fitting the He et al. [2010] model, as well as a list of which parameters were fixed or estimated. We again note that the parameter μ here represents death rate minus the net immigration rate of the population. Unlike in the deterministic and stochastic SEIR models with school-term forcing, the values of the incubation and recovery rates were estimated for this model. Additionally, the basic reproduction number was estimated using the mean transmission rate and (2.14).

3.7.3   Model strengths and weaknesses

As a discrete state stochastic model, the He et al. [2010] can exhibit both the endemic dynamics of big cities (including annual and biennial dynamics) and the episodic dynamics in smaller cities. If the immigration term ι is greater than zero, the model can allow for stochastic fadouts and reemergence of the disease. A weakness of this model is that as a continuous time Markov chain model, it is more computationally expensive to simulate and fit to data. It is more computationally expensive than the stochastic SEIR model discussed earlier due to its multiple sources of stochasticity (multiplicative noise in the transmission process, stochasticity from the Euler mul-

Table 3.4: Descriptions and values of parameters for the He et al. [2010] model at the MLE

Symbols	Description	Range	Value	Fixed/Estimated
μ	Death minus net immigration rate	$[0, \infty)$	0.01 yr^{-1}	Estimated
σ	Incubation rate	$[0, \infty)$	41.21 yr^{-1}	Estimated
γ	Recovery rate	$[0, \infty)$	62.69 yr^{-1}	Estimated
R_0	Basic reproduction number	$[0, \infty)$	30.45	Estimated
a	Amplitude of seasonality	$[0, 1]$	0.12	Estimated
τ	Recruitment delay	$[0, \infty)$	4 yrs	Fixed
ρ	Reporting probability	$[0, 1]$	0.50	Estimated
ψ	Reporting overdispersion	$[0, \infty)$	0.05	Estimated
α	Mixing exponent	$[0, \infty)$	0.98	Estimated
σ_{SE}	White-noise intensity on $S \to E$ transition rate	$[0, \infty)$	0.0700 yr$^{1/2}$	Estimated
ι	Mean number of visiting infectives	$[0, \infty)$	4.5876	Estimated
c	Cohort entry fraction	$[0, 1]$	1.0000	Estimated
$N(0)$	Initial population size	$[0, \infty)$	2,445,368	Fixed
$S(0)/N(0)$	Initial fraction of susceptibles	$[0, 1]$	0.03	Estimated
$E(0)/N(0)$	Initial fraction of exposed	$[0, 1]$	5.17×10^{-5}	Estimated
$I(0)/N(0)$	Initial fraction of infected	$[0, 1]$	5.14×10^{-5}	Estimated
$R(0)/N(0)$	Initial fraction of recovered	$[0, 1]$	0.97	Estimated

tiinomial draws, and stochasticity from the measurement model). While the many sources of noise should help improve the likelihood of this model, this may also lead to larger variation in predictions

3.7.4 Summary of the model

- The model is a continuous time Markov chain model with discrete states and transitions given by (3.37)–(3.47) and (3.48)–(3.51).

- Cases can be aggregated over any time period by adding the number of tran-

sitions from I to R during that time period.

- Report are assumed to have an overdispersed binomial distribution approximated by equation (3.7) with constant reporting probability ρ.

- Seasonally changing transmission rates assumed to be due to school-term forcing and given by (3.6).

- Actual London birth rates and population were used as covariates.

- Entry into the susceptible class is assumed to have a delay τ of about four years corresponding to the delay in school entry.

- The cohort effect (3.40) adds an additional source of seasonal forcing in the model.

- The model includes stochasticity in the transitions which allows for more variability in dynamics.

- Discrete states allow for stochastic fadeouts in small populations.

- If $\iota > 0$ this parameter allows for infections from outside the region.

- Of the four focus models, this model is the most computationally expensive to simulate and fit.

3.8 Overall summary

In this chapter, different models of pre-vaccine era measles transmission were presented. A review of the history of measles transmission models was presented including the papers by Bartlett [1957, 1960] which established the critical community size of measles, and London and Yorke [1973] which presented the necessity of using changing transmission rates to reflect school-term forcing. The work of Schenzle

[1984] on the RAS model, Keeling and Grenfell [1997] on the PRAS model was also discussed, as well as well as the contributions of Finkenstadt et al. [1998] on the effect of birth rates on periodicity, and Cauchemez and Ferguson [2007] on using a continuous time Markov chain model for measles were also discussed.

An exposition of the work done for this project was given in section 3.2. The pre-vaccine era London data set used to fit and compare the different measles models, as well as the transient and vaccine era data set that was not used for fitting but used to compare the models, were described in section 3.3. Four measles models were chosen to be studied in detail for this thesis: (1) the deterministic SEIR model with school-term forcing, (2) the stochastic SEIR model with school-term forcing, (3) the Bjornstad et al. [2002] TSIR model, and (4) the He et al. [2010] model. A detailed description and discussion of these four models were given in the four subsequent sections. Finally, in Table 3.5 we present a summary of the four models with their important features and characteristics.

In the next chapter, we compare the four focus models at their MLEs from fitting to the pre-vaccine era London data. We compare them using likelihood, AIC and residuals. We also perform a preliminary study of how these models at their pre-vaccine era MLEs extend to vaccine era.

Table 3.5: Summary comparison of the four focus models.

Specification	Deterministic model with school-term forcing	Stochastic Model with school-term forcing	Bjornstad et al. [2002] model	He et al. [2010] model
Process model form	ODE system	Continuous time Markov chain	Discrete time Markov chain	Continuous time Markov chain
Sources of seasonal forcing	School-term forcing	School-term forcing	General time-varying transmission rate	School-term forcing and cohort effect
Sources of stochasticity	Measurement model only	Measurement and process model	Measurement and process model (includes additive noise)	Measurement and process model (includes multiplicative noise in contact rates)
Sources of delays	Delay in school entry (about 4 years)	Delay in school entry (about 4 years)	Delay due to maternal immunity (about 16 weeks)	Delay in school entry (about 4 years)
Measurement model	Over-dispersed binomial (3.7)	Over-dispersed binomial (3.7)	Over-dispersed binomial (3.36)	Over-dispersed binomial (3.7)
Assumptions on contact	Mass action	Mass action	Allows for variation from mass action using a mixing parameter	Allows for variation from mass action using a mixing parameter
Model Strengths	Can generate regular endemic dynamics. Simple formulation and fast parameter fitting.	Can generate regular endemic dynamics and fadeouts which lead to disease extinction.	Can generate regular endemic and episodic dynamics, relatively fast parameter fitting	Can generate regular endemic and episodic dynamics such as stochastic fadeouts
Model weaknesses	Cannot exhibit episodic dynamics, no variability in state processes	Cannot fully exhibit episodic dynamics due to no disease immigration term, computationally expensive.	No latent period, strict assumption on time scales that disease generation time is comparable to time between observations.	Computationally expensive.

4

COMPARISON OF THE MODELS

In this chapter, the four focus models from chapter 3 are evaluated with regards to their performance in the pre-vaccine, transient and vaccine era. Recall that these four models were (1) the deterministic SEIR model with school-term forcing, (2) the stochastic SEIR model with school-term forcing, (3) the Bjornstad et al. [2002] TSIR model, and (4) the He et al. [2010] model.

In section 4.1, the performance of four focus models in the pre-vaccine era are compared using likelihood, Akaike information criterion (AIC) and residuals. In section 4.2, a comparison of the performance of these four models into the transient and vaccine eras is presented. This comparison was done by using a simple extension of the models fitted to pre-vaccine era data and incorporating the historical vaccine coverage estimates from the UK. Finally, in section 4.3, sample simulations of the models at their best estimated parameters are presented. The general features of these simulations, such as their periodicity and heights of epidemics are discussed and compared to the data.

4.1 Pre-vaccine era comparison

4.1.1 AIC

As described in section 2.7, the AIC is one way of selecting the best among competing models. This criterion selects the best model by balancing bias and flexibility as measured by the log-likelihood and the number of model free parameters respectively. The best among the competing models is the one with the smallest AIC value. These values are presented in Table 4.1. The table shows the estimated log-likelihood, the number of parameters, and the AIC score for each model. We also present the likelihood of the best log-SARMA $(1,1) \times (1,1)_{26}$ model fitted to the data. SARMA stands for Seasonal Auto-Regressive Moving Average and $y_{1:N}$ is said to be log-SARMA if $\log(y_{1:N}) = \log y_1, \ldots, \log y_N$ is SARMA (He et al. [2009]). Thus the log-SARMA model is a non-mechanistic empirical model. The likelihood for the log-SARMA model shown in Table 4.1 is obtained by using *auto.arima* function of the forecast package of CRAN (Hyndman et al. [2017]) on $\log y_{1:N}$, adjusted by subtracting $\sum_{n=1}^{N} \log y_n$ to transform the likelihood back from the logarithmic scale (He et al. [2009]).

From Table 4.1, it is clear that the best among these models using the AIC criterion is the He et al. [2010] model with an AIC of 6656. This is the only model in this study that performs better that the empirical log-SARMA model (AIC of 6884) with regards to this criterion. This is followed by the stochastic SEIR model with school-term forcing (AIC of 8150), then the deterministic model with school-term forcing (AIC of 8670). The least preferable model here happens to be the Bjornstad et al. [2002] TSIR model (AIC of 8966). The ranking by AIC is strongly dictated by the ranking of the models by likelihood. Adjusting by the number of parameters to calculate the AIC yielded the same ordering of the models as the ordering we would have derived from ranking the models using likelihood only.

We note that it is perhaps not surprising that the Bjornstad et al. [2002] model has the worst likelihood since, as discussed in the previous chapter, the parameters of the process model were not fitted using maximum likelihood with the given reporting model. Because of this we present another method of model comparison in the next section, one that does not use likelihood.

Table 4.1: Log-likelihood and AIC's of the four focus models at their best parameter estimates.

Model	No. of parameters	Log-likelihood	AIC
Deterministic model with school-term forcing	5	-4330	8670
Stochastic model with school-term forcing	5	-4070	8150
Bjornstad et al. [2002] model	28	−4455	8966
He et al. [2010] model	15	-3313	6656
SARMA model	6	-3436	6884

4.1.2 Residuals

The four focus models were also compared with respect to their residuals at their fitted parameter estimates. This is a commonly used, direct and very intuitive way to compare the goodness of fit of the different models. However, it is important to remember that the parameters of these models were not determined by minimizing the residuals (least squares fitting). Applying the standard least squares procedure would have been equivalent to maximizing likelihood under the assumption of a Gaussian measurement model with constant variance regardless of the value of the number of measles cases. In contrast, we had instead assumed over-dispersed binomial measurement model which allows for more variance during the peaks of epidemics than during the troughs.

The sum of square residuals (SSR) is the standard residual used for considering model fit to data. Here we also consider the sum of square relative residuals (SSRR). The SSR is a measure of the absolute deviation of the model from the data, and the SSRR is a measure of the deviation of the model from the data relative to the magnitude of the recorded cases. Given data on measles reports denoted by reports$_{1:n}$ and model simulation output given by $y_{1:n}$, the SSR and SSRR are defined as follows:

$$SSR = \sum_i^n \Big(\text{reports}_i - y_i\Big)^2$$

$$SSRR = \sum_i^n \Big(\frac{\text{reports}_i - y_i}{\text{reports}_i}\Big)^2.$$

In the SSR, all residuals are equally weighted. Large deviations from recorded cases are penalized irrespective of the magnitude the recorded cases, unlike how the overdispersed binomial which would be more forgiving when there are large numbers of cases as opposed to smaller number of cases.

The SSRR is more lenient to large deviations that are associated with large observations. The mean and variation of the SSR and SSRR of the four focus models at their MLEs are presented in Table 4.2. These were calculated by looking at the distribution of the SSR and SSRR for each model using 500 simulations.

Table 4.2: Residual statistics of the four focus models at their best parameter estimates in the pre-vaccine era.

Model	Mean SSR	SSR S.E.	Mean SSRR	SSRR S.E.
Deterministic model with school-term forcing	9.574×10^8	9.849×10^7	8.643×10^3	2.529×10^3
Stochastic model with school-term forcing	9.318×10^8	2.088×10^8	1.757×10^4	1.465×10^4
Bjornstad et al. [2002] model	6.949×10^8	1.606×10^8	1.522×10^3	2.805×10^3
He et al. [2010] model	1.395×10^9	3.782×10^8	1.455×10^5	1.891×10^5

With regard to the SSR, the Bjornstad et al. [2002] model is now by far the best performing model with an SSR of 6.949×10^8. This is followed by the stochastic model with school-term forcing and the deterministic model with school-term forcing, which are comparable in terms of SSR. The He et al. [2010] model, which was the best model in terms of likelihood, is the worst performing model here. This may be due to the model having the most inherent variability and noise.

The best model according to the SSRR is still the Bjornstad et al. [2002] model with the least SSRR (1.522×10^3). Up next, the deterministic model and stochastic model with school-term forcing switch places, with the deterministic model with school-term forcing having a slightly smaller SSRR (i.e. 8.643×10^3) as compared to the stochastic model with school-term forcing SSRR of 1.757×10^4. Additionally, the stochastic model with school-term forcing had more variability in the SSRR statistic. This suggests that the deterministic model does better than the stochastic model at the troughs of epidemics. The He et al. [2010] model is again the lowest ranked model by this criterion with an SSRR of 1.891×10^5.

Overall, the Bjornstad et al. [2002] model performs the best with regard to the residual statistics. This is followed by the deterministic model with school-term forcing and the stochastic model with school-term forcing which are comparable. Curiously, the He et al. [2010] model, which was the best model according to likelihood and AIC, is the lowest ranked with respect to the residual statistics in the pre-vaccine era. This result is a matter of interest for future work on the connections between noise and likelihood.

4.2 Comparison of simple extensions of the model into the transient and vaccine eras

We also compared how simple extensions of the four focus models would perform in the transient and pre-vaccine era. We did this by changing the vaccine coverage $v(t)$ from zero in the pre-vaccine era to values from the historical measles vaccine coverage in England and Wales which is given in Table 4.3. We did not fit the models to the transient and vaccine era data, we just used the four focus models fitted to the pre-vaccine era and considered how the models performed in the new period by looking at the SSR and SSRR, as well as by looking at the qualitative features of the dynamics. The residuals for the transient and vaccine era are presented in Table 4.4.

Table 4.3: Historical vaccine coverage in England and Wales obtained from The National Archives of Public Health England at `http://webarchive.nationalarchives.gov.uk/20140714110743/http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/VaccineCoverageAndCOVER/EpidemiologicalData/coverVaccineUptakeData/`.

Year	Vaccine coverage
1968	0.33
1969	0.46
1970	0.51
1971	0.53
1972	0.52
1973	0.46
1974	0.46
1975	0.48
1976	0.48
1977	0.51
1978	0.53
1979	0.55
1980	0.58
1981	0.60
1982	0.63
1983	0.68
1984	0.71

Table 4.4: Residual statistics of the four focus models in the vaccine era

Model	SSR	SSR S.E.	SSRR	SSRR S.E.
Deterministic model with school-term forcing	1.258×10^9	2.337×10^8	1.605×10^5	5.431×10^4
Stochastic Model with school-term forcing	3.088×10^8	1.225×10^8	3.551×10^4	2.762×10^4
Bjornstad et al. [2002] model	5.715×10^8	9.054×10^7	9.971×10^4	4.221×10^4
He et al. model model	3.635×10^8	1.597×10^8	7.899×10^4	1.036×10^4

With regard to the SSR in the transient and vaccine eras, the stochastic model with school-term forcing is the best performing model with an SSR of 3.088×10^8. This is followed by the He et al. [2010] with 3.635×10^8, and then Bjornstad et al. [2002] model with 5.715×10^8. The deterministic model with school-term forcing is the worst performing model here with an SSR of 1.258×10^9. In the next section, we show that the bad performance of the extension of the deterministic model is due to the use of continuous states not allowing for fadeouts of the disease and yielding periodic large epidemics in the vaccine era.

The ranking of the extensions of the models does not change with regard SSRR. The stochastic model with school-term forcing still performs the best with 3.551×10^4. This is followed by the He et al. [2010] model (7.899×10^4), then the Bjornstad et al. [2002] model with 9.971×10^4. The model with the worst SSRR (i.e. 1.605×10^5) in the transient and vaccine era is the deterministic model with school-term forcing.

Overall, the stochastic model with school-term forcing performs the best with regard to the residual statistics in the transient and vaccine era. This is followed by the He et al. [2010] model, then the Bjornstad et al. [2002] model. The worst performing model here is the deterministic model with school-term forcing.

4.3 Model simulations

Deterministic model with school-term forcing

Sample simulations (including measurement noise) of the deterministic model with school-term forcing at its MLE is shown in Figure 4.1. The transition from regular annual to biennial epidemics of measles are evident in the model trajectory, however the model trajectory appears to be a little late in getting the transition right, and the peaks from the biennial epidemics are not as high. The model does well with the troughs of the epidemics, explaining its relatively small SSRR.

This model is clearly exhibiting multi-year periodic epidemics in the vaccine era, something that we do not observe in the actual data. This is due to the model having continuous states which does not allow for the $I(t)$ state to go to zero. Instead, $I(t)$ goes to very low values which allow for the number of susceptibles to build up until this pool becomes so large it is able to set off a large epidemic. This clear discrepancy between one of the model's key features and the data indicates that this model is very inadequate at modeling the vaccine era dynamics.
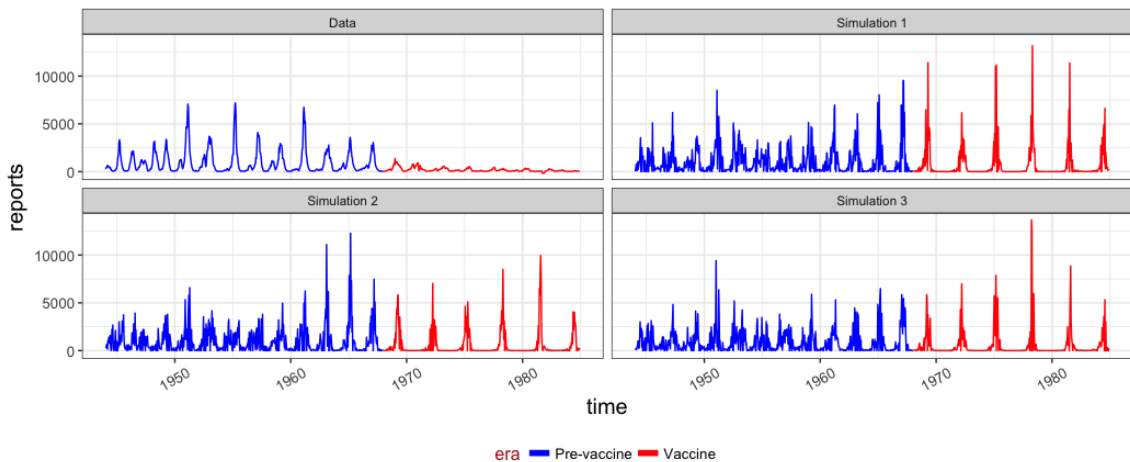

Figure 4.1: Simulations of the deterministic SEIR Model with school-term forcing and its simple extension in the vaccine era.

Stochastic SEIR model with school-term forcing

Sample simulations of the stochastic SEIR model with school-term forcing at its MLE are shown in Figure 4.2. The simulations exhibit more stochasticity than the simulations of the deterministic model in Figure 4.1. The model also exhibits the expected transition from annual to biennial epidemics in the pre-vaccine era, however the added stochasticity in the model's state process makes the transition less clear.

Unlike the deterministic model, the stochastic model has discrete states and it is therefore possible to have extinction of the disease when prevalence has gone down in the vaccine era. This means the stochastic model is less likely to decisively exhibit multi-year periodic epidemics in the vaccine era that the deterministic model displays. However it also means that the disease has a significant probability of going extinct as we see in the sample simulations. While extinction yields better SSR and SSRR than the large vaccine era epidemics of the deterministic model, this still indicates that the stochastic model that we use is inadequate for the vaccine era. Inclusion of infection from outside the region (such as the ι term in the He et al. [2010] model) would help prevent the disease from exhibiting permanent extinction and would lead to better simulations in the vaccine era.
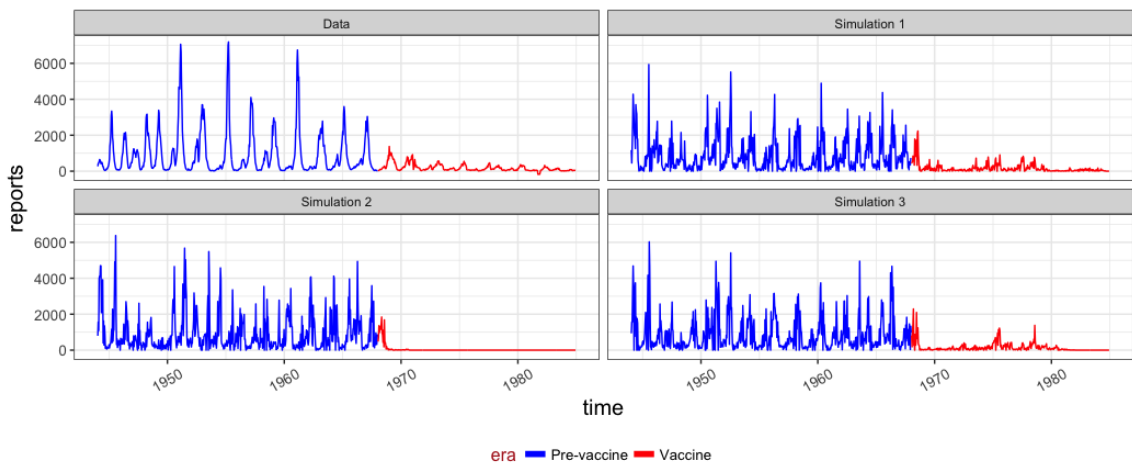

Figure 4.2: Comparison of the data and simulations of the stochastic model with school-term forcing extended into the vaccine era.

Bjornstad et al. [2002] model

Figure 4.3 presents sample simulations of the Bjornstad et al. [2002] model at its best parameter estimates. Here it is clear that the model can exhibits the biennial epidemics of measles, but the annual epidemics are less clear. This models also appears to have less noise than the stochastic SEIR model due to being a discrete time process.

The TSIR model exhibits large deviations from data mostly in the biennial periodic dynamics, where there are larger peaks. This may be one reason for its relatively small SSRR.

The transient and vaccine era dynamics of the TSIR model is clearly different form what we see in the data. The number of simulated reports given by the model is much higher in magnitude, and it also exhibits a transient periodic epidemics of about every 2–3 years before settling to annual epidemics that are still larger than what we see in the data. It is clear that the simple extension of this model is not sufficient for modeling the transient and vaccine era dynamics in London.
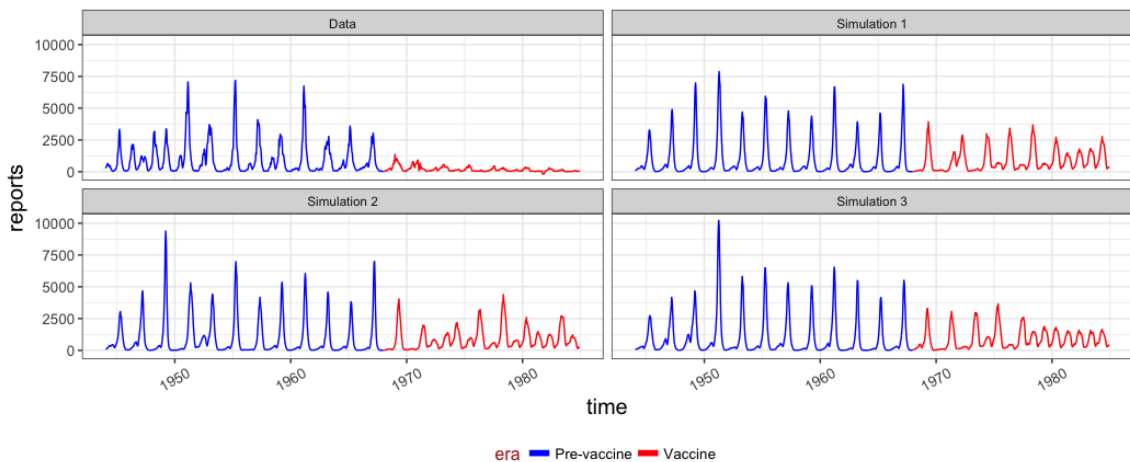

Figure 4.3: Comparison of the data and simulations of the Bjornstad et al. [2002] model extended into the vaccine era.

He et al. [2010] model

The simulations of the He et al. [2010] model at its MLE is compared to the data in Figure 4.4. Here we can see why this model, which has the best likelihood, is the lowest ranked in terms of SSR and SSRR in the pre-vaccine era. Though individual simulations of its measles cases clearly exhibit the early pre-vaccine era annual dynamics and late pre-vaccine era biennial dynamics, there is larger variation in the timing of the transition as well as the occurrence of large epidemics in the simulations.

The He et al. [2010] model simulations into the vaccine era had worse SSR and SSRR than the stochastic SEIR model with school-term forcing. However, as already mentioned, the stochastic SEIR model obtains this score by often driving the disease to extinction and not allowing for reintroduction of the disease. The He et al. [2010] model, on the other hand, yields irregular epidemics in the transient and vaccine era. This still does not quite look like the data, but at least there are no clear regular large epidemics (like the deterministic SEIR model) or extinctions (like the stochastic SEIR model). However, in many cases the simulations still exhibited occasional big epidemics with peaks that are much higher than what can be found in the data.
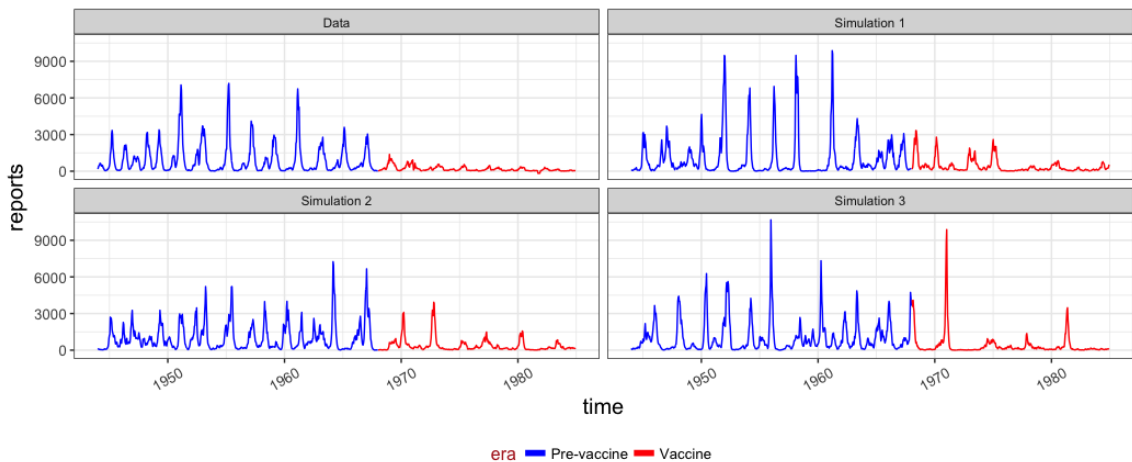

Figure 4.4: Comparison of the data and simulations of the He et al. [2010] model extended into the vaccine era.

4.4 Summary

In this chapter we compared the four focus models of this thesis which were discussed in detail in chapter 3. The models were compared with regards to how well they fit the data in the pre-vaccine era (the data used to fit the models), and the transient and vaccine era (the data that was not used to fit the model).

In section 4.1, the models were compared in the pre-vaccine era using their likelihoods and criteria such as AIC and residuals. The He et al. [2010] model ended up performing the best under the AIC and likelihood, but the Bjornstad et al. [2002] model performed the best in terms of SSR and SSRR in the pre-vaccine era (refer to Tables 4.1 and 4.2). The He et al. [2010] model performed the worst in terms of both the SSR and SSRR, perhaps due to the presence of more noise in this model. This case of having the best likelihood but the worst residuals is is a rather surprising result that merits further study.

In section 4.2, the performance of these models into the transient and vaccine eras were compared. This was done by extending the four models using the pre-vaccine era's estimated parameters and including estimates of the vaccine coverage. The residual statistics in the vaccine era were presented in Table 4.4. Here the stochastic model with school-term forcing performed the best by having a high fraction of its simulation exhibiting extinction of the disease. As a result of all this we conclude that none of the simple extensions of these four focus models perform well in the vaccine era.

In section 4.3 the qualitative features of the simulations of the four models were compared to the features we see in the data. Most models can exhibit some rough transition from annual to biennial dynamics in the pre-vaccine era, however they do not always exhibit the transition at the right time.

The simulation of the models extended into the vaccine era were also compared. After the start of mass vaccination, all of the simulations from all models ex-

hibited a decrease in measles reports. However, aside from this, none of the models actually did very well in capturing the transient and vaccine era dynamics which consisted of much smaller regular epidemics. The deterministic SEIR with school-term forcing model instead yielded simulations in the vaccine era that had too large epidemics with long interepidemic periods. The stochastic SEIR model with school-term forcing did not have this feature, and though some simulations look a little like the data, many other simulations exhibited disease extinction which was not be found in the data. The Bjornstad et al. [2002] model simulations into the transient and vaccine era showed transient periodic epidemics that were 2–3 years apart which later on gives way to smaller annual epidemics that still had too high peaks relative to the data. The He et al. [2010] model simulations showed irregular epidemics that had peaks that were higher (sometimes much higher) than what can be found in the data.

In the next chapter, we conclude this thesis by summarizing the motivation and goals for this project, the results, open questions and future directions on how the project can be further extended.

5

CONCLUSION

The transmission dynamics of measles during the pre-vaccine era is a well-studied topic in mathematical epidemiology. Many models have been formulated to explain the dynamics of the disease. These models were derived under various assumptions and come with different levels sophistication. The goal of this project was to present an overview of the history of measles models and to carefully compare the assumptions and performance of some well-known models. In particular, we considered the following four models based on the standard susceptible-exposed-infected-recovered (SEIR) model framework:

1. Deterministic SEIR model with school-term forcing

2. Stochastic SEIR model with school-term forcing

3. Bjornstad et al. [2002] time series SIR (TSIR) model with a general time-dependent transmission rate

4. He et al. [2010] model, a stochastic SEIR model with school-term forcing, cohort effect and possibility of infection from outside the population

A summary of the important features of these four models is presented in Table 3.5 at the end of Chapter 3.

5.1 Summary of results

In this thesis we carefully reproduced the analysis of each of the four focus models and fitted them to bi-weekly data from 1944–1964 (in the pre-vaccine era) in London. The performance of each model was evaluated using the likelihood and model comparison criteria such as AIC and residuals. Additionally, we performed a preliminary comparison of the performance of these models into the transient and vaccine eras by extending the fitted pre-vaccine era models using known estimates of the vaccine coverage.

The He et al. [2010] model performed the best among all models with respect to both AIC and likelihood. This continuous time, discrete state model which included school-term forcing, the cohort effect and possibility of infections arising from outside the population (if $\iota > 0$) was also the only one of the four models to beat the non-mechanistic SARMA model in terms of AIC. Unfortunately the He et al. [2010] model is also the most computationally expensive to simulate of the four focus models. The remaining three models are ranked in the following order: The stochastic SEIR model with school-term forcing had the next highest AIC, followed by the deterministic SEIR model with school-term forcing. The least preferred model using likelihood and the AIC criterion was the Bjornstad et al. [2002] model. We again note that this may not be surprising since this model was not fitted using maximum likelihood estimation.

Interestingly, the Bjornstad et al. [2002] model performed the best in the residual statistics (both SSR and SSRR) than the other three models in the pre-vaccine era. And even more interestingly, the He et al. [2010] performed the worst. We hypothesize that this is perhaps due to the presence of more noise in the He et al. [2010] model. This unexpected result will be an interesting topic for future study

The models (with parameters fitted to the pre-vaccine era data) were extended into the transient and vaccine era by taking into account the historical national

vaccine coverage data from the UK. Here it is the extension of the stochastic SEIR model with school-term forcing performed the best with respect to the SSR and SSRR in the transient and vaccine era. This was accomplished by having a significant fraction of its simulations exhibiting extinction of the disease with no possibility of reintroduction. However, we do not expect that this is the actually what is happening with measles.

None of the model simulations into the transient and vaccine eras seem to exhibit the patterns we see in the recorded data. After the start of mass vaccination, all of the simulations from all the models exhibited a decrease in measles reports, however the other features of the data were not captured. The deterministic SEIR with school-term forcing model simulations displayed too large epidemics with inter-epidemic periods that were too long. The stochastic SEIR model with school-term forcing exhibited too much disease extinction. The Bjornstad et al. [2002] model had transient periodic epidemics that had too long periods then annual epidemics, all of which had too high peaks. The He et al. [2010] model simulations had irregular epidemics that can sometimes have too high peaks. From this, and from the fact that the stochastic SEIR performed best in terms of the residuals (by letting the disease go to extinction often), we concluded that none of the four focus models, even the He et al. [2010] model which included many important mechanisms for measles transmission, can simply be extended into the vaccine era. We note that all of these models are homogeneous models of measles transmission. Perhaps structured population models can lead to better fits in the vaccine era.

5.2   Limitations of the study

In this study we compared measles models by looking at how well these models fit to the pre-vaccine era data from London, a large city in the UK. However, it has been

observed that pre-vaccine era measles in the UK can often be classified into one of two large classes of dynamics: endemic dynamics with clear periodicity exhibited in large cities, and the episodic outbreaks presented in smaller communities (Bjornstad et al. [2002]). Since we focused on the city of London in this thesis we do not have a careful comparison of how these models would perform in smaller communities. Our expectation is that the deterministic SEIR model would perform worse than the other three models, just as it performed worst in the transient and vaccine eras, since the continuous state model would be unable to produce stochastic "fadeouts" we expect to see in smaller communities.

Additionally, we performed a preliminary comparison of how these models would perform when extended to the vaccine era by incorporating estimates of the vaccine coverage. However, the vaccine era data was not used for fitting the models. Thus it is possible that the ranking of how well the models fit the vaccine era data may change if the vaccine era data was also used for fitting the models.

5.3   Recommendations for future work

As we have already mentioned, the curious result of the He et al. [2010] model having the best likelihood and AIC, but worst SSR and SSRR (which are more intuitive measures of model fit) in the pre-vaccine era merits further investigation. Additionally, to obtain a more comprehensive review of the different measles models, the models can be fitted to more of the available data from other UK cities. In particular they should be fitted to at least one smaller city that exhibits episodic rather than endemic measles dynamics. This will help in comparing the generalizability of the dynamics of the models.

A more exhaustive review of how the models perform in the transient and vaccine era could also be performed. It is notable that simple extensions of the models to

the vaccine era do not appear to be able to recreate these dynamics. We note that all of the four focus models we considered assume that the population exhibits homogeneous contact rates. The results of this study indicate that homogeneous models appear to work well for the pre-vaccine era dynamics. However, it is widely accepted that population contact rates may be structured by age and that this may be relevant for disease transmission models (Schenzle [1984]). Perhaps age-structured contact rates may play a more important role in modeling the transition from pre-vaccine era to the vaccine era.

Bibliography

Hirotugu Akaike. A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, dec 1974. ISSN 15582523. doi: 10.1109/TAC.1974.1100705. URL `http://ieeexplore.ieee.org/document/1100705/`.

Hirotugu Akaike. Likelihood of a model and information criteria. *Journal of Econometrics*, 16(1):3–14, 1981. ISSN 03044076. doi: 10.1016/0304-4076(81)90071-3.

R M Anderson and R M May. Regulation and stability of host-parasite population interactions. I. Regulatory processes. *Journal of Animal Ecology*, 47:219–247, 1978.

R M Anderson and R M May. No Title. *Infectious Diseases of Humans: Dynamics and Control*, 1991.

M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, 2002. ISSN 1053587X. doi: 10.1109/78.978374. URL `http://ieeexplore.ieee.org/document/978374/`.

T. Barrett. The molecular biology of the morbillivirus (measles) group. *Biochemical Society Symposia*, 53:25–37, 1987. cited By 10.

M S Bartlett. Measles Periodicity and Community Size. *Journal of the Royal Sta-*

tistical Society. Series A (General), 120(1):48–70, 1957. ISSN 00359238. doi: 10.2307/2342553. URL `http://www.jstor.org/stable/2342553`.

M. S. Bartlett. The Critical Community Size for Measles in the United States. *Journal of the Royal Statistical Society. Series A (General)*, 123(1):37, 1960. ISSN 00359238. doi: 10.2307/2343186. URL `http://www.jstor.org/stable/10.2307/2343186?origin=crossref`.

Alexander D Becker and Sinead E Morris. Title An Implementation of the TSIR Model. *Comprehensive R Archive Network*, 2017. doi: 10.1111/1467-9876.00187>. URL `https://cran.r-project.org/web/packages/tsiR/tsiR.pdf`.

Ottar N Bjornstad, Barbel F Finkenstadt, and Bryan T Grenfell. Dynamics of Measles Epidemics: Estimation Scaling Of Transmission Rates Using A Time Series SIR Model. *The Ecological Monographs*, 2002.

Francis L Black. *Viral Infections of Humans*. Springer US, Boston, MA, 1984 edition, 1984. ISBN 978-1-4684-4729-3. doi: 10.1007/978-1-4684-4727-9. URL `http://link.springer.com/10.1007/978-1-4684-4727-9`.

Carles Bretó, Daihai He, Edward L. Ionides, and Aaron A. King. Time series analysis via mechanistic models. *The Annals of Applied Statistics*, 3(1):319–348, mar 2009. ISSN 1932-6157. doi: 10.1214/08-AOAS201. URL `http://projecteuclid.org/euclid.aoas/1239888373`.

S. Cauchemez, A.-J. Valleron, P.-Y. Boëlle, A. Flahault, and N.M. Ferguson. Estimating the impact of school closure on influenza transmission from sentinel data. *Nature*, 452(7188):750–754, 2008. doi: 10.1038/nature06732. cited By 285.

Simon Cauchemez and Neil M Ferguson. Likelihood-based estimation of continous-time epidemic models from time-series data: application to measles transmission in London. *The Royal Society*, 2007.

Centers for Disease Control. Measles | Signs and Symptoms | CDC, 2017. URL `https://www.cdc.gov/measles/about/signs-symptoms.html`.

Centers for Disease Control and Prevention. *Epidemiology and Prevention of Vaccine-Preventable Diseases.* Public Health Foundation, Washington D.C., 13th edition, 2015. URL `http://bookstore.phf.org/.http://www.cdc.gov/vaccines/pubs/pinkbook/index.html`.

A Cliff, P Haggett, and M Smallman-Raynor. No Title. *Measles: An Historical Geography of a Major Human Viral Disease from Global Expansion to Local Retreat, 1840-1990*, 1993.

Ciro A de Quadros. Can measles be eradicated globally? *Bulletin of World Health Organization*, 2004.

Richard C Dicker. *Principles of Epidemiology in Public Health Practice.* Number May. Center for Disease Control, 2006. URL `https://www.cdc.gov/ophss/csels/dsepd/ss1978/SS1978.pdf`.

O Diekmann, J a P Heesterbeek, and M G Roberts. The construction of next-generation matrices for compartmental epidemic models. *Journal of the Royal Society, Interface / the Royal Society*, 7(47):873–885, 2010. ISSN 1742-5689. doi: 10.1098/rsif.2009.0386.

Arnaud. Doucet, Nando. De Freitas, and Neil Gordon. *Sequential Monte Carlo methods in practice.* Springer, 2001. ISBN 9780387951461.

David J D Earn, Pejman Rohani, Benjamin M Bolker, and Bryan T Grenfell. A Simple Model for Complex Dynamical Transitions in Epidemics. *Science*, 287 (5453):667–670, 2000a. ISSN 0036-8075, 1095-9203. doi: 10.1126/science.287.5453.667.

D.J.D. Earn, P. Rohani, B.M. Bolker, and B.T. Grenfell. A simple model for complex dynamical transitions in epidemics. *Science*, 287(5453):667–670, 2000b. doi: 10.1126/science.287.5453.667. cited By 369.

W J Edmunds, N J Gay, M Kretzschmar, R G Pebody, H Wachmann, and ESEN Project. European Sero-epidemiology Network. The pre-vaccination epidemiology of measles, mumps and rubella in Europe: implications for modelling studies. *Epidemiology and infection*, 125(3):635–50, dec 2000. ISSN 0950-2688. URL `http://www.ncbi.nlm.nih.gov/pubmed/11218214http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2869647`.

S Ellner, A R Gallant, and J Theiler. Detecting nonlinearity and chaos in epidemic data. *Epidemic Models: Their Structure and Relation to Data*, pages 229–247, 1995.

S.P. Ellner, B.A. Bailey, G.V. Bobashev, A.R. Gallant, B.T. Grenfell, and D.W. Nychka. Noise and nonlinearity in measles epidemics: Combining mechanistic and statistical approaches to population modeling. *American Naturalist*, 151(5): 425–440, 1998. doi: 10.1086/286130. cited By 95.

B. Finkenstadt, M. Keeling, and B. Grenfell. Patterns of density dependence in measles dynamics. *Proceedings of the Royal Society B: Biological Sciences*, 265(1398):753–762, may 1998. ISSN 0962-8452. doi: 10.1098/rspb.1998.0357. URL `http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.1998.0357`.

B.F. Finkenstädt, Bryan T B.T. Grenfell, Barbel F Finkenstadt, and Bryan T B.T. Grenfell. Time series modelling of childhood diseases: a dynamical systems approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 49(2):187–205, 2000. ISSN 00359254. doi: 10.1111/1467-9876.00187. URL `http://doi.wiley.com/10.1111/1467-9876.00187`.

R A Fisher. On the Mathematical Foundations of Theoretical Statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 222(594-604):309 LP – 368, jan 1922. URL `http://rsta.royalsocietypublishing.org/content/222/594-604/309.abstract`.

B. Grenfell and J. Harwood. (meta) population dynamics of infectious diseases. *Trends in Ecology and Evolution*, 12(10):395–399, 1997. doi: 10.1016/S0169-5347(97)01174-9. cited By 221.

B T Grenfell and A P Dobson. Ecology of infectious diseases in natural populations. *Ecology of Infectious Diseases in Natural Populations*, 1995.

Bryan T Grenfell, Ottar N Bjørnstad, Bärbel F Finkenstädt, Source Ecological Monographs, and No May. Model DYNAMICS OF MEASLES EPIDEMICS : SCALING NOISE , DETERMINISM , AND PREDICTABILITY WITH THE TSIR MODEL. *Ecological Monographs*, 72(2):185–202, 2014.

B.T. Grenfell, O.N. Bjørnstad, and B.F. Finkenstädt. Dynamics of measles epidemics: Scaling noise, determinism, and predictability with the tsir model. *Ecological Monographs*, 72(2):185–202, 2002. cited By 124.

Daihai He, Edward L Ionides, and Aaron A King. Supplement to " Plug-and-play inference for disease dynamics: measles in large and small towns as a case study ". *Journal of The Royal Society Interface*, 7(43), 2009. URL `http://rsif.royalsocietypublishing.org/content/royinterface/suppl/2009/06/10/rsif.2009.0151.DC1/rsif20090151supp1.pdf`.

Daihai He, Edward L Ionides, and Aaron A King. Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *The Royal Society*, 2010.

Herbert W. Hethcote. Qualitative analyses of communicable disease models. *Mathematical Biosciences*, 28(3-4):335–356, 1976. ISSN 00255564. doi: 10.1016/0025-5564(76)90132-2. URL `http://people.kzoo.edu/barth/math280/articles/communicable{_}disease.pdf`.

Herbert W Hethcote. The Mathematics of Infectious Diseases. *Society for Industrial and Applied Mathematics*, 2000.

Peter J. Hotez. Texas and Its Measles Epidemics. *PLOS Medicine*, 13(10):e1002153, oct 2016. ISSN 1549-1676. doi: 10.1371/journal.pmed.1002153. URL `http://dx.plos.org/10.1371/journal.pmed.1002153`.

Rob Hyndman, Mitchell O'Hara-Wild, Christoph Bergmeir, Slava Razbash, and Earo Wang. Title Forecasting Functions for Time Series and Linear Models. *CRAN*, 2017.

Aaron A. King Ionides, Dao Nguyen, and Edward L. Statistical Inference for Partially Observed Markov Processes via the {R} Package {pomp}. *Journal of Statistical Software*, 69(12):1 —- 43, 2016. doi: 10.18637/jss.v069.i12.

E L Ionides, C Bretó, and A A King. Inference for nonlinear dynamical systems. *Proceedings of the National Academy of Sciences of the United States of America*, 103(49):18438–43, dec 2006. ISSN 0027-8424. doi: 10.1073/pnas.0603181103. URL `http://www.ncbi.nlm.nih.gov/pubmed/17121996http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3020138`.

Edward L Ionides. Discussion of "Feature Matching in Time Series Modeling" by Y. Xia and H. Tong. *Statist. Sci.*, 26(1):49–52, 2011. doi: 10.1214/11-STS345C. URL `http://dx.doi.org/10.1214/11-STS345C`.

Edward L Ionides, Dao Nguyen, Yves Atchadé, Stilian Stoev, and Aaron A King. Inference for dynamic and latent variable models via iterated, perturbed

Bayes maps. *Proceedings of the National Academy of Sciences of the United States of America*, 112(3):719–24, jan 2015. ISSN 1091-6490. doi: 10.1073/ pnas.1410597112. URL `http://www.ncbi.nlm.nih.gov/pubmed/25568084http: //www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4311819`.

John A Jacquez and Carl P Simon. Qualitative Theory of Compartmental Systems. *SIAM Review*, 35(1):43–79, mar 1993. ISSN 0036-1445. doi: 10.1137/1035003. URL `http://epubs.siam.org/doi/10.1137/1035003`.

M J Keeling and B T Grenfell. Disease extinction and community size: modeling the persistence of measles. *Science (New York, N.Y.)*, 275(5296):65–7, jan 1997. ISSN 0036-8075. URL `http://www.ncbi.nlm.nih.gov/pubmed/8974392`.

B E Kendall, W M Schaffer, L F Olsen, C W Tidd, and B L Jorgensen. Using chaos to understand biological dynamics. *Predictability and Nonlinear Modelling in Natural Sciences and Economics*, pages 184–203, 1994.

W. O. Kermack and A. G. McKendrick. Contributions to the Mathematical Theory of Epidemics. III. Further Studies of the Problem of Endemicity. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 141 (843), 1933. URL `http://rspa.royalsocietypublishing.org/content/141/ 843/94`.

WO O. Kermack and A. G. AG McKendrick. Contributions to the Mathematical Theory of Epidemics. II. The Problem of Endemicity. *Proc. R. Soc. Lond.*, 138(834):55–83, 1932. ISSN 1364-5021. doi: 10.1098/rspa.1932. 0171. URL `http://rspa.royalsocietypublishing.org/content/138/834/ 55http://www.math.utah.edu/{~}bkohler/Journalclub/kermack1932.pdf`.

A. Aaron King. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License., 2016a. URL

https://creativecommons.org/choose/results-one?license{_}code=
by-nc{&}jurisdiction={&}version=4.0{&}lang=en.

Aaron A. King. Model-based Inference in Ecology and Epidemiology, 2016b. URL
http://kingaa.github.io/short-course/.

Aaron A. King, Dao Nguyen, and Edward L. Ionides. Statistical Inference for Partially Observed Markov Processes via the R Package pomp. *Journal of Statistical Software*, 69(12):1–43, 2016. ISSN 15487660. doi: 10.18637/jss.v069.i12. URL http://www.jstatsoft.org/v69/i12/http://arxiv.org/abs/1509.00503.

Aaron A King, Edward L Ionides, Carles Martinez Bretó, Stephen P Ellner, Matthew J Ferrari, Bruce E Kendall, Michael Lavine, Dao Nguyen, Daniel C Reuman, Helen Wearing, and Simon N Wood. *pomp: Statistical Inference for Partially Observed Markov Processes*, 2017. URL https://cran.r-project.org/web/packages/pomp/http://kingaa.github.io/pomp.

Genshiro Kitagawa. Non-Gaussian State—Space Modeling of Nonstationary Time Series. *Journal of the American Statistical Association*, 82(400):1032–1041, dec 1987. ISSN 0162-1459. doi: 10.1080/01621459.1987.10478534. URL http://www.tandfonline.com/doi/abs/10.1080/01621459.1987.10478534.

Michael Y. Li and James S. Muldowney. Global stability for the SEIR model in epidemiology. *Mathematical Biosciences*, 125(2):155–164, feb 1995. ISSN 00255564. doi: 10.1016/0025-5564(95)92756-5. URL http://linkinghub.elsevier.com/retrieve/pii/0025556495927565.

Wayne P. London and James A. Yorke. RECURRENT OUTBREAKS OF MEASLES, CHICKENPOX AND MUMPS. *American Journal of Epidemiology*, 98(6):453–468, dec 1973. ISSN 1476-6256. doi: 10.1093/oxfordjournals.

aje.a121575. URL `https://academic.oup.com/aje/article-lookup/doi/10.1093/oxfordjournals.aje.a121575`.

Richard K. Miller and Anthony N. Michel. *Ordinary differential equations.* Academic Press, 1982. ISBN 9780124972803. URL `http://www.sciencedirect.com/science/book/9780124972803`.

Webb Miller, Vanessa M Hayes, Aakrosh Ratan, Desiree C Petersen, Nicola E Wittekindt, Jason Miller, Brian Walenz, James Knight, Ji Qi, Fangqing Zhao, Qingyu Wang, Oscar C Bedoya-Reina, Neerja Katiyar, Lynn P Tomsho, Lindsay McClellan Kasson, Rae-Anne Hardie, Paula Woodbridge, Elizabeth A Tindall, Mads Frost Bertelsen, Dale Dixon, Stephen Pyecroft, Kristofer M Helgen, Arthur M Lesk, Thomas H Pringle, Nick Patterson, Yu Zhang, Alexandre Kreiss, Gregory M Woods, Menna E Jones, and Stephan C Schuster. Genetic diversity and population structure of the endangered marsupial Sarcophilus harrisii (Tasmanian devil). *Proceedings of the National Academy of Sciences of the United States of America*, 108(30):12348–53, jul 2011. ISSN 1091-6490. doi: 10.1073/pnas.1102838108. URL `http://www.ncbi.nlm.nih.gov/pubmed/21709235http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3145710`.

Rachael Milwid, Andreea Steriu, Julien Arino, Jane Heffernan, Ayaz Hyder, Dena Schanzer, Emma Gardner, Margaret Haworth-Brockman, Harpa Isfeld-Kiely, Joanne M Langley, and Seyed M Moghadas. Toward Standardizing a Lexicon of Infectious Disease Modeling Terms. *Frontiers in public health*, 4(SEP):213, 2016. ISSN 2296-2565. doi: 10.3389/fpubh.2016.00213. URL `http://www.ncbi.nlm.nih.gov/pubmed/27734014`.

Sharmistha Mishra, David N Fisman, and Marie-Claude Boily. The ABC of terms used in mathematical models of infectious diseases. *J Epidemiol Community Health*, 2010.

D. Mollison and S. Ud Din. Deterministic and stochastic models for the seasonal variability of measles transmission. *Mathematical Biosciences*, 117(1-2):155–177, 1993. doi: 10.1016/0025-5564(93)90021-2. cited By 14.

Alexander Morton and Bärbel F. Finkenstädt. Discrete time modelling of disease incidence time series by using Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, 54(3):575–594, jun 2005. ISSN 00359254. doi: 10.1111/j.1467-9876.2005.05366.x. URL `http://doi.wiley.com/10.1111/j.1467-9876.2005.05366.x`.

L.F. Olsen and W.M. Schaffer. Chaos versus noisy periodicity: Alternative hypotheses for childhood epidemics. *Science*, 249(4968):499–504, 1990. cited By 186.

D.A. Rand and H.B. Wilson. Chaotic stochasticity: A ubiquitous source of unpredictability in epidemics. *Proceedings of the Royal Society B: Biological Sciences*, 246(1316):179–184, 1991. cited By 148.

D. Schenzle. An age-structured model of pre- and post-vaccination measles transmission. *IMA Journal of Mathematics Applied in Medicine and Biology*, 1(2):169–191, 1984. cited By 140.

I.B. Schwartz. Multiple stable recurrent outbreaks and predictability in seasonally forced nonlinear epidemic models. *Journal of Mathematical Biology*, 21(3):347–361, 1985. doi: 10.1007/BF00276232. cited By 75.

G. Sugihara, B. Grenfell, and R.M. May. Distinguishing error from chaos in ecological time series. *Philosophical Transactions - Royal Society of London, B*, 330(1257):235–251, 1990. cited By 102.

Gerald Etapelong Sume, André Arsène Bita Fouda, Marie Kobela, Salomé Nguelé, Irène Emah, Peter Atem, Daddy Mbida, and Kondé Njock. Epidemiology and clinical characteristics of the Measles outbreak in the Ny-

lon Health District, Douala Cameroon: a retrospective descriptive cross sectional study. *The Pan African medical journal*, 13:66, 2012. ISSN 1937-8688. URL `http://www.ncbi.nlm.nih.gov/pubmed/23346280http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3549450`.

P. van den Driessche and J. Watmough. Further notes on the basic reproduction number. In F. Brauer, P. van den Driessche, and J. Wu, editors, *Mathematical Epidemiology*, chapter 6, pages 159–178. Springer, 2008.

B. G. Williams and C. Dye. Infectious disease persistence when transmission varies seasonally. *Mathematical Biosciences*, 145(1):77–88, 1997. ISSN 00255564. doi: 10.1016/S0025-5564(97)00039-4.