Stochastic Modelling of Measles in a Disease-Eliminated Setting in Europe

Student

10 September, 2017

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Mauris semper vel libero eget maximus. Sed non nibh ac orci pretium placerat. Cras vulputate dictum libero nec vulputate. Suspendisse pharetra eu nibh sit amet rhoncus. Pellentesque lacus odio, ultrices aliquam dictum ut, sagittis ac neque. Cras consectetur sem nulla, quis consectetur nunc vulputate id. Vestibulum non ullamcorper nisi, eu tincidunt dui. Curabitur ut suscipit sapien. Nam at dui rutrum ligula tristique ullamcorper id sed tortor. Curabitur nec sem vitae sapien feugiat ultrices eu et eros. Duis sit amet facilisis ligula. Donec semper augue magna, ut hendrerit lorem maximus eget. Cras quis hendrerit sem.

Contents

1 Acknoledgements

I would like to thank Dr. Sebastian Funk, to whom I owe a great deal for his patience and guidance during this project.

I would also like to thank my family, and partner, Annie, who all provided tremendous support throughout the process.

In loving memory of my late grandmothers, Faye Boyer and Cynthia Browning, both of whom passed away during my studies in London.
Through your love, I pursue my dreams.

2 Introduction

Measles is one of 25 vaccine preventable illnesses identified by the WHO as a global eradication target in their Global Vaccine Action Plan. While elimination efforts in some areas of the globe have seen great success, Europe still bears a disproportionate burden of the disease. This is due, at least in part, to low vaccination coverage in several nations[1].

While data on effective vaccination rates vary based on national surveillance capacity, regional efforts, like the European Sero-Epidemiology Network (ESEN), have provided more robust measurements ranging over a decade. Their findings reveal that all nations with an endemic spread of the disease fall short of the herd immunity threshold (around 96%)[2]. This shortfall reflects the recent trend of hesitancy on the part of parents to immunize their child with both doses of the MMR vaccine required for full coverage.

Beyond the immediate effects of endemic measles transmission within a country, it is reasonable to question whether these nations also put the elimination efforts of their neighbors at risk. It is this question that this project seeks to answer, particularly as more nations start to reach their elimination targets and are declared "measles eliminated" by the European divsion of the WHO[3].

3 Aims and Objectives

The primary aim is to explore the susceptibility of measles-eliminated countries in Europe to experience an epidemic following new case importation. There are two ideas behind this aim:

Firstly, I reason that there is some interplay between population size, the rate of new case introduction, and effective vaccination rate which affects it's susceptibility to an epidemic;

Secondly, I reason that there exists some period of time, after which "measles eliminated" countries would again be susceptible to large scale outbreaks, dependent on the interplay descibed in the former premise.

To accomplish this, I first created a stochastic transmission model of measles in the country. Using population data from the United Nations and seroprevalence data from the World Health Organization, I extrapolated initial population values for the compartmental model.

Beyond this, I created a modelling framework in the R programming language in order to handle the insertion of infected individuals into the model and determine the impact after running several iterations. As the

4 Materials and Methods

4.1 Coding

All model functions were coded in R[4] with C++ suplement. The utilized packages were as follows:

- adaptivetau[5]
- ggplot2[6]
- data.table[7]
- Rcpp[8]
- Foreach[9]
- microbenchmark[10]
- profvis[11]

To increase modelling performance, the Foreach package parallelized the modelling function to utilize all availible processor threads. The data.table package provided the "data.table" R object that is capable of much faster data sorting and subsetting methods than the standard packages. Rcpp served as a linker package between R and C++, allowing for computationally intensive portions of the R code to be offloaded to C++ and called within the R environment. Finally, the microbenchmark and profvis packages were essential for code optimization and profiling.

In addition to the model functional code, this document was created in LaTeX and compiled to PDF using the XeLaTeX compiler.

4.2 SEIR Compartmental Model

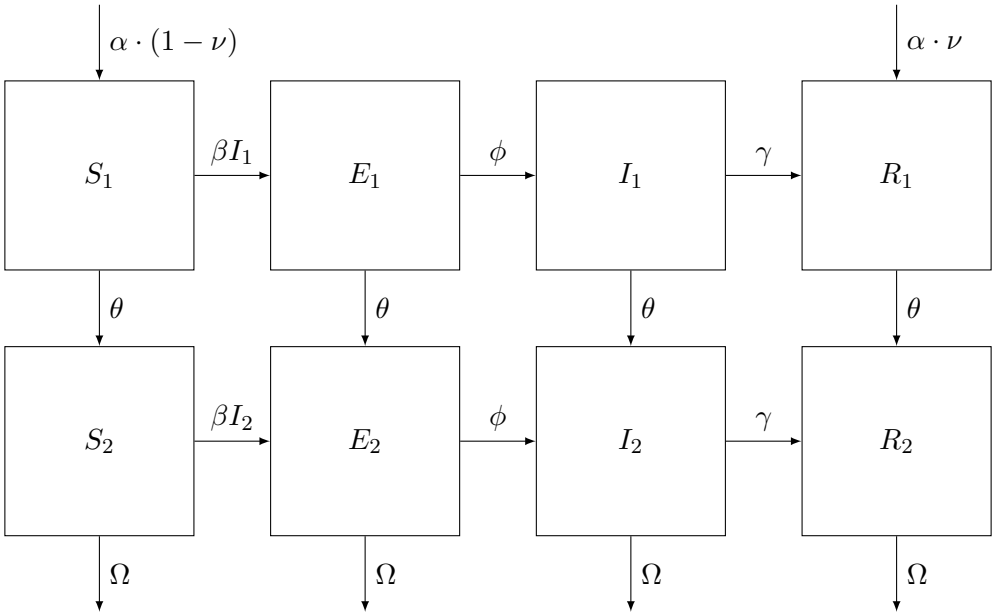

Figure 1: Age stratified SEIR Model with vital dyanmics

Where:

β_n = Rate of contact between infectious and susceptible persons
ϕ = Rate of onset of infectiousness subsequent to being infected
γ = Rate of recovery from measles from infectious period
α = Crude birth rate
ν = Effective vaccination rate at birth
Ω = Crude death rate
θ = Rate of aging from young to old compartments
f_t = New case introduction rate

To explore the infection dynamics of measles, a SEIR model was selected with two age compartments, subdividing younger and older populations. This was chosen for a few reasons. Firstly, it allows for the modelling of heterogeneous mixing within these age groups. Additionally, it allows for the model to more closely align with the age strata reported in seroprevalence data, thereby reducing the number of assumptions which would need to be made to fit incoming data. Finally, the selection of two compartments over some higher dimension limits the mathematical complexity of both modeling the data as well as mapping the parameter space while also allowing for some variation in age demographics.

4.3 Model Dynamics as Deterministic ODEs

The dynamics which describe the compartmental model can be written deterministically as ordinary differential equations :

$$\frac{\mathrm{d}S_1}{\mathrm{d}t} = N\alpha\,(1-\nu) - S_1\,(\beta I_1 + \theta) \tag{1}$$

$$\frac{\mathrm{d}E_1}{\mathrm{d}t} = S_1\beta I_1 - E_1\,(\phi + \theta) \tag{2}$$

$$\frac{\mathrm{d}I_1}{\mathrm{d}t} = E_1\phi - I_1\,(\gamma + \theta) \tag{3}$$

$$\frac{\mathrm{d}R_1}{\mathrm{d}t} = N\alpha\nu + I_1\gamma - R_1\theta \tag{4}$$

$$\frac{\mathrm{d}S_2}{\mathrm{d}t} = S_1\theta - S_2\,(\beta I_2 + \Omega) \tag{5}$$

$$\frac{\mathrm{d}E_2}{\mathrm{d}t} = E_1\theta + S_2\beta I_2 - E_2\,(\phi + \Omega) \tag{6}$$

$$\frac{\mathrm{d}I_2}{\mathrm{d}t} = I_1\theta + E_2\phi - I_2\,(\gamma + \Omega) \tag{7}$$

$$\frac{\mathrm{d}R_2}{\mathrm{d}t} = R_1\theta + I_2\gamma - R_2\Omega \tag{8}$$

4.4 The Model as a Stochastic System

Since the overall model is stochastic, we can represent it as a continuous Markov-chain process like so: This diagram serves to highlight the possible transitions when the system is simulated.

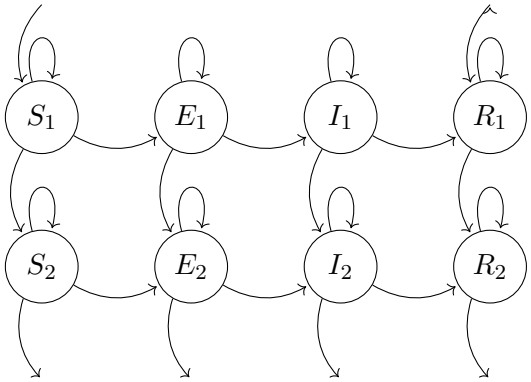

Figure 2: Markov-chain representation of the SEIR Model

The ODEs still describe the average propensity for each transition in the same way as the compartmental SEIR model in Figure 1.

The method used to solve the stochastic system is an explicit τ-leaping variation of the Gillespie Algorithm. In R, the package which provides this functionality is called adaptivetau. The deterministic differential equations are supplied and evaluated by adaptivetau as propensity functions. In turn, it applies those functions to the list of state space transitions between passed as a vector.

Gillespie describes the mechanism as it pertains to chemical physics in his first publication on the topic[12]. Let us define the propensity of any transition in the model as a_j. The goal of the algorithm is then to find the value of τ which results in the smallest appreciable increase in a_j of a given transition.

Cao et al. extend this by providing an algorithm for efficiently selecting values of τ [13]. In more plain terms, the "explicit τ-leaping" method depends on bounding the change in event rate of each transition at a given time by the transition rate function which produces the largest value in that same period. The calculated rates are also a product of an initial tolerance value, which Cao suggests should be set at 0.03. Evaluating all transition event rates with this method produces a set of possible values of τ, the minimum of which is the optimal value.

With the value of τ calculated, the probability density function (PDF) of the event rate for any state space transition E_j in that time step is Poisson-distributed:

$$P\left(E_j\right) \sim Poisson\left(a_j \cdot \tau\right) \tag{9}$$

As a consequence, for sufficiently high transition propensities a_j, or values of τ, the PDF of the

event rate described in Equation 9 will follow the normal distribution :

$$P\left(E_j\right) \sim N\left(a_j \cdot \tau, \sqrt{a_j \cdot \tau}\right) \tag{10}$$

This process provides an approximation of the output expected from the exact algorithm by maximizing the time-to-event between data points while minimizing the rate of change of any transition. As a result, the output is dense only where the rate function is sufficiently high (as it is at the beginning of an outbreak). Because it is an approximate stochastic simulation, there is an inevitable loss of resolution over a fixed-τ approach. However, this comes with a reduced computation time as the simulation does not require nearly the same amount of data points to approximate the trajectory of the simulation as an exact method would.

The mechanism of case introduction into the model was designed as a Markov-jump process (referred to as m_t) to allow for a delayed introduction event. Beyond this, the jump process also has an upper cutoff which roughly limits the period where measles infected individuals can enter the model. Doing so ensures that the simulation always runs longer than the length of the outbreak. The dynamics of the jump process is defined in this way :

$$m(t) = \begin{cases} \epsilon, & \text{for } m_{start} < t \leq m_{end} \\ 0, & \text{for } m_{start} > t, \\ & \text{or } t > m_{end} \end{cases} \tag{11}$$

where $\epsilon =$ the rate of new measles case introduction, $m_{start} =$ the start time of new importation, and $m_{end} =$ the end time of new importation

4.5 Estimating Model Parameters

There were three nations chosen for analysis using the model mentioned above: Sweden, Malta, and Latvia. All three have been declared measles-free by the WHO and see few annual cases. Despite this, the populations of these three countries have entirely different seroprevalence values with Latvia having the highest proportion of seronegative individuals and Sweden having the lowest. The table below highlights this variation along with nationally-reported MCV1 and MCV2 rates.

	MCV1							MCV2						
	2016	2015	2014	2013	2012	2011	2010	2016	2015	2014	2013	2012	2011	2010
Sweden	97%	98%	97%	97%	97%	96%	97%	95%	95%	95%	95%	95%	95%	94%
Malta	93%	89%	98%	99%	93%	84%	73%	86%	91%	94%	88%	91%	85%	97%
Latvia	93%	96%	95%	96%	90%	92%	95%	89%	92%	89%	92%	92%	92%	93%

Table 1: MCV1 and MCV2 coverage rates as reported to the WHO [14]

Population demographic data for the project was acquired from the United Nations Department of Economic and Social Affairs, Population Division (UN DESA). Measles seroprevalence data was made available by findings from ESEN2 presented to the World Health Organization.

As previously mentioned, the model will have two age compartments in an effort to reduce the number of dimensions. However, the WHO report provides seroprevalence data in several age categories, and the appropriate proportion of measles seronegativity in the younger age compartment must be extrapolated. It stands to mention that many of the data are quite dated. Several countries, such as Sweden, have not submitted updated data in over 20 years, therefore some liberty was taken in best estimating these population values. Where possible, sensitivity analyses of the parameters were conducted to quantify the impact on model results.

4.6 Calculating R0

We can calculate R0 by taking the leading eigenvalue of the Next Generation Matrix. Probably not important

4.7 Defining Model Steady-State Conditions

This section may not need to exist

4.8 Visualizing Hyperparameter Space

To determine the smallest inputs into the model which would result in an epidemic, I analyzed the hyperparameter space of two model parameters. I employed whole grid search to avoid the posibility of missing the global optimum, which approximation methods such as random search may do without appropriate parameter weighting.

Let the two parameter inputs be defined as:

$$x \in \{1, 2, \cdots, n\} \tag{12}$$

and,

$$y \in \{1, 2, \cdots, n\} \tag{13}$$

where:
$x =$ number of new cases of measles added to the model, and
$y =$ number of "insertion events" which occur during one model simulation.

The cost function, $f(x, y)$, provides scalar output of the largest outbreak length when evaluating the model with x and y as inputs.

The three dimensional surface of the hyperparameter space is found by evaluating $f(x, y)$ with all reasonable combinations of x and y. The three values can then be expressed as cartesian coordinates and plotted. Importantly, this will help to reveal local maxima, minimizing for the input values.

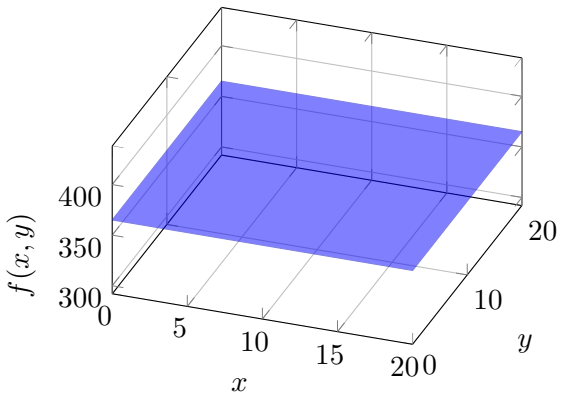

Figure 3: Example surface plot; maxima threshold shown at $f(x, y) = 365$

5 Results

There will be some results here

6 Discussion

There will be discussion of results here.

7 Recommendations

There will be some conclusions drawn here from the data. Hopefully they will be useful.

References

[1] World Health Organization (WHO). Measles vaccines: WHO position paper April 2017. *Weekly Epidemiological Record*, 92(17):205228, Apr 2017.

[2] Nick Andrews, Annedore Tischer, Annette Siedler, and Richard G Pebody. Towards elimination: measles susceptibility in Australia and 17 European countries. *Bulletin of the World Health Organization*, 86, Mar 2008.

[3] David H Sniadack, Natasha S Crowcroft, David N Durrheim, and Paul A Rota. Roadmap to elimination-standard measles and rubella surveillance. *Weekly Epidemiological Record*, 92(9/10):97105, Mar 2017.

[4] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017.

[5] Philip Johnson. *adaptivetau: Tau-Leaping Stochastic Simulation*, 2016. R package version 2.2-1.

[6] Hadley Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York, 2009.

[7] Matt Dowle and Arun Srinivasan. *data.table: Extension of 'data.frame'*, 2017. R package version 1.10.4.

[8] Dirk Eddelbuettel and Romain François. Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, 40(8):1–18, 2011.

[9] Revolution Analytics and Steve Weston. *foreach: Provides Foreach Looping Construct for R*, 2015. R package version 1.4.3.

[10] Olaf Mersmann. *microbenchmark: Accurate Timing Functions*, 2015. R package version 1.4-2.1.

[11] Winston Chang and Javier Luraschi. *profvis: Interactive Visualizations for Profiling R Code*, 2017. R package version 0.3.3.

[12] Daniel T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *The Journal of Chemical Physics*, 115(4):1716–1733, 2001.

[13] Cao Yang, Daniel T Gillespie, and Linda R Petzold. Adaptive explicit-implicit tau-leaping method with automatic tau selection. *The Journal of Chemical Physics*, 126(22), Jun 2007.

[14] World Health Organization (WHO). Who vaccine-preventable diseases: monitoring system, 2017.