Project, FM1015 Modelling of Dynamic Systems

Bernt Lie

September 18, 2020

1 Introduction

1.1 General instructions

- This project is a *work requirement* in course FM1015 Modelling of Dynamic Systems:
  - Contributing in the project and the project report is a *work requirement*, and a 50% score is required for the project report to pass this work requirement.
  - Contributing in an oral presentation of the group project is a *work requirement*, and good presentation and understanding of the project report is a requirement to pass the work requirement.
  - The project problem will constitute the *core* of the individual exam in November/December, 2020.[1]
- You are expected to work in a group of four (± 1); you should establish the groups yourselves. The *group* is expected to write and submit a single group report in a *Canvas group folder*:
  - You are recommended to form groups with a mixture of background (EET, EPE, IIA, PT; different nationalities, etc.) — this will give the group a good variation in process knowledge, programming background, and communication practice.
  - Those who have not established groups by September 18, 2020, will be placed in groups by me.
  - Do *not* establish a Canvas group folder yourself — instead, send an e-mail to the lecturer, in Canvas, where you state who will be group members, and then the lecturer will establish a *Canvas group folder*.[2]
  - A mid-way report should be submitted in the Canvas group folder by Friday, October 23, 2020 at 10:00. In the mid-way report, a dynamic model with first simulation results should be included.

[1] There may be questions unrelated to the project problem. If so, these will constitute a minor part of the exam.

[2] **Reason**: A "group room" established by students allows students to be member in multiple groups, and it is impossible to keep track of whether students are in a project group or not. The lecturer can establish a type of "group room" where students can only be member of one group.

  - A single document final report, PDF format, maximum of 15 pages + a single cover page, is to be submitted in the Canvas group folder by Friday, November 13, 2020 at 23:59.[3]

- The oral project presentation:
  - Each group has 7 minutes for a presentation (PowerPoint, etc.), and every group member needs to present something.
  - Following the presentation will be a brief examination of the work, where every group member will be asked at least one question. In total, 15 minutes will be allotted to each group (presentation + questions).
  - *Campus students* will give the presentation in class on Friday November xx, 2020, from 10:15 until finish; *industry master students* are recommended to present in class on this day.
  - *On-line students* who (normally) can not be present, will give their presentation via MS Teams at a different time — *individual appointments* for Teams presentations will be set up with the lecturer.

The notation used in this project more or less follows the standard notation in the course.

1.2 Background

The spread of the virus *SARS-CoV-2*[4], and subsequent disease *covid-19*[5], has had a strong impact on public health and economy in the world in 2020. In order to understand the spread of this infectious disease, and be able to reduce the negative impact of it, a large amount of data has been made available[6], and many models have been proposed in order to predict the effect of actions[7] to reduce the spread.

Pandemics such as *covid-19* are not new. The (bacterial) bubonic plague[8] that came to Europe ca. 1346-47 became known as the *Black Death* (1346–1353), and killed an estimated 60% of the population in Europe (50-80 million), and some 75–200 million worldwide; see Fig. 1 for a graphic description of its spread in Europe. Poland closed its borders, and was more or less spared from the plague; Iceland was partially spared since sailors died before the ships arrived to Iceland.

As an example, the Black Death came to Norway by ships to the Hansa city Bjørgvin (currently: Bergen) in 1349, spread rapidly, and after several waves of the plague throughout the 1350s had killed one third of the population. It took some 400 years for the population level to return to the pre-Black Death level. This first wave of the bubonic plague inspired literature [9]; later waves inspired theology[10] (Germany ca. 1527), and Sir Isaac Newton did some of his important work while self quarantining due to the bubonic

[3] Yes, maximum 15+1 pages. Not a single page extra for computer code, plots, etc.

[4] Severe acute respiratory syndrome (SARS) coronavirus (CoV) 2

[5] COrona VIrus Disease (20)19

[6] See, e.g., https://www.worldometers.info/coronavirus/ for some data.

[7] See, e.g., https://covid19.healthdata.org/united-states-of-america

[8] https://en.wikipedia.org/wiki/Bubonic_plague

[9] E.g., Giovanni Boccaccio: The Decameron, ca. 1353, and Geoffrey Chaucer: The Canterbury Tales, ca. 1387-1400.

[10] E.g., https://witness.lcms.org/2020/read-luthers-letter-about-plague/

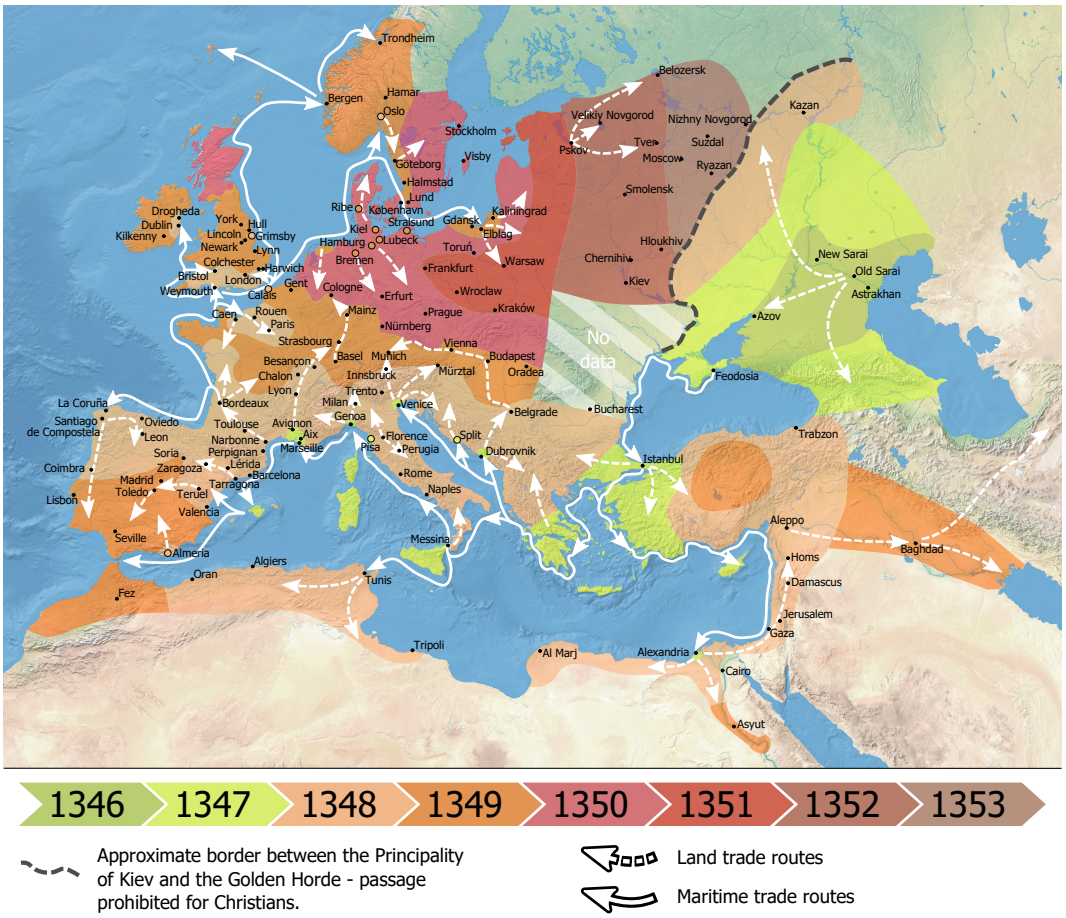

Figure 1: The origin and early spread of the Black Death in Italy: first evidence of plague victims from 14th-century Liguria (northern Italy). The map is based on work by historian O.J. Benedictow, and is provided to the Natural Earth public domain map data set (https://www.naturalearthdata.com/) by Flappiefh; see https://no.wikipedia.org/wiki/Svartedauden#/media/Fil:1346-1353_spread_of_the_Black_Death_in_Europe_map.svg.

plague in Oxford (1665–1667). Occasional incidences with the bubonic plague still occur, e.g., a couple of minor outbreak in China in 2020. Today, the bubonic plague can be treated using antibiotics.

The *1918 flu pandemic* ("The Spanish flu", which most likely originated in Kansas, USA) from 1918-1920 was an unusually deadly influenza pandemic caused by the *H1N1 influenza A* virus, and swept the world in 3(4) waves while infecting some 500 million people (a third of the world's population then), with death toll in the range 17-50 million people — perhaps as much as 100 million people. In Europe, the death tolls were in the range 0.4–1.3% of the population; in India, some 6.1% of the population died, while in some indigenous populations, up to 80–90% of the population died (Labrador, Canada; some places in Alaska).

The decade following the 1918 flu pandemic saw important progress in infectious disease modeling which still inspires such models. Typically, the population of N individuals was partitioned into a deterministic 3 compartmental model consisting of those who are susceptible to an infection (S), those who have been infected (I), and those who have recovered (R), Kermack & McKendrick (1927), so-called SIR models. Approximately at the same time, the stochastic *Reed-Frost* model was formulated which describes the same SIR types, Schwabe et al. (1977).

With the current covid-19 pandemic, it is of interest to use scientific methods and models to understand how such diseases spread. Such models can be developed and analyzed using the knowledge acquired in course FM1015 Modelling of Dynamic Systems.

1.3 Project aim

This project is designed to give experience in formulating and solving dynamic models of systems, and thus to aid in the learning of the content of course FM1015. Furthermore, the purpose of the project and the presentation is to give the students experience in group work and communication of findings. Finally, because the written exam will be based on the project work, the project functions as preparation for the written exam.

More specifically, in this group project, infectious disease models are to be formulated using classical deterministic number/material balance principles in combination with ideas of "chemical" reactions. The resulting models are to be analyzed wrt. infection spread, and the effects of mitigating policies.

A first step in the project is to get familiar with the simplest infection model (the SIR mode) and simulation of a simple case of measles infection. A *second step* involves formulating and implementing a dynamic model of covid-19 infection from the literature, and use model parameters valid for Italy. A *third step* involves expanding the covid-19 model with a more proper mitigation response model. A *fourth step* involves combining the covid-19 model for Italy and Spain with equal exchange of travelers, to see how various infection levels mix. A *fifth step* involves studying the possibilities of quenching the epidemics by reaching herd immunity, or by using vaccination.

At the end of this project, you will be able to:

1. understand the basic principles of developing a model for infections,
2. solve the thermal model using a computer tool (OpenModelica, Python, Julia, MATLAB, etc.), and
3. understand how models can be used to study mitigation policies (control strategies)

.

2 Epidemiology models

2.1 Introduction

> "Elegant ladies, as I believe you know, the wisdom we mortals possess does not merely consist of remembering things past and apprehending the present, but on the basis of these two activities being able to predict the future, which is considered by serious men to be the highest form of human intelligence."
>
> Boccaccio, Giovanni (1349-1351/-52). *The Decameron*, p. 851. Translated by Wayne A. Rebhorn. W. W. Norton & Company. Kindle Edition.

Historically, the population of N individuals was partitioned into a deterministic 3 compartmental model consisting of those who are susceptible to an infection (S), those who have been infected (I), and those who have recovered (R), Kermack & McKendrick (1927). Approximately at the same time, the stochastic *Reed-Frost* model was formulated which describes the same SIR types; this model was published much later, Schwabe et al. (1977). Classical epidemiological models use a *spatial description* (Euler description) where the focus is on a number of compartments of fixed area/people with fixed attributes (age, immune level, etc.) where people move in and out of the various compartments/attributes. An alternative could be a *material description* (Lagrangian description) where one instead track the status of each individual or a group of individuals of similar attributes, and how these move about in the world, get into proximity with other people, and get infected with a certain probability. Such *material descriptions* are sometimes referred to as *agent-based* models.[11] Today, stochastic descriptions are in common use, based on the assumptions of Poisson distributed events. These stochastic descriptions are either formulated as continuous stochastic differential equations, or as event-driven, discrete models.

In this paper, an introduction to both classical deterministic models, as well as stochastic models are given.[12] However, the project *tasks* for the students is limited to deterministic, differential equation models.

2.2 Number balance

In epidemiology models, individuals are categorized into n_{X} types/classes, e.g., class X_j with $j \in \{1, \ldots, n_{\mathrm{X}}\}$. The number of individuals in the *population* will be noted N, and the number of individuals in *class/compartment* X_j will be denoted X_j. The *fraction* (per capita) of individuals in class X_j will be denoted $\check{X}_j$, normally $\check{X}_j \triangleq \frac{X_j}{N}$.[13]

The flow rate of individuals will be denoted by $\dot{N}_j$, e.g., $\dot{X}_j$ for the flow rate of individuals in class X_j. Such a flow rate may be *influent* flow rate $\dot{X}_j^{\mathrm{i}}$ (immigration), *effluent* flow rate $\dot{X}_j^{\mathrm{e}}$ (emigration), or *generation* flow rate due to some transformative mechanisms, $\dot{X}_j^{\mathrm{g}}$.

[11] Some presentations of epidemiology refer to Lagrangian and Eulerian *movement* in an unconventional way, e.g., Martcheva (2015), pp. 389–392.

[12] In reality, even more complex models are used by the public health authorities, https://www.fhi.no/sv/smittsomme-sykdommer/corona/koronavirus-modellering/

[13] A notation more in line with the notation of the course *lecture notes* would be to use N_{X_j} for the number of individuals in class X_j and x_{X_j} for the per capita/fraction (= "mole fraction"). However, this makes the notation somewhat convoluted, so here we use the simpler notation of X_j and $\check{X}_j$ — to minimize the use of subscripts.

To simplify the notation, we can stack up the number of individuals X_j into vector $X \in \mathbb{N}_0^{n_\mathrm{X}}$ and similarly the rates $\dot{X}_j^\mathrm{i}$ into vector $\dot{X}_\mathrm{i} \in \mathbb{Z}^{n_\mathrm{X}}$, rates $\dot{X}_j^\mathrm{e}$ into vector $\dot{X}_\mathrm{e} \in \mathbb{Z}^{n_\mathrm{X}}$, and rates $\dot{X}_j^\mathrm{g}$ into a vector $\dot{X}_\mathrm{g} \in \mathbb{Z}^{n_\mathrm{X}}$.[14]

A number n_r of transformative reaction mechanisms may take place; each class X_j may be influenced by several reaction mechanisms. In general, such mechanisms are related to the classes of individuals by a *stoichiometric matrix* $\nu \in \mathbb{Q}^{n_\mathrm{r} \times n_\mathrm{X}}$ which has one row per reaction mechanism, and one column per class n_X.[15] With r_k denoting the *rate* of reaction mechanism $k \in \{1, 2, \ldots, n_\mathrm{r}\}$, these rates can be stacked into vector $r \in \mathbb{Q}^{n_\mathrm{r}}$. Next, we introduce the vector *generation rate* per capita, $r_\mathrm{g} \in \mathbb{Q}^{n_\mathrm{X}}$, defined by

$$\dot{X}_\mathrm{g} \triangleq N r_\mathrm{g}$$

where N is the total population number. The generation rate per capita, r_g, is then related to the reaction mechanism rate r as

$$r_\mathrm{g} = \nu^\mathsf{T} r.$$

It follows that element j of r_g is given as $r_j^\mathrm{g} = \left(\nu^\mathsf{T} r\right)_j$.

Associated with the rate of reaction mechanism r_j is a time constant τ_j which indicates some mean time for this transformation, alternatively an associated frequency factor $k_j = 1/\tau_j$.

We can now formulate a number balance for the individuals in the compartments,

$$X\left(t + \Delta t\right) = X\left(t\right) + \int_t^{t+\Delta t} \left(\dot{X}_\mathrm{i}\left(\theta\right) - \dot{X}_\mathrm{e}\left(\theta\right) + N\left(\theta\right) \cdot \left(\nu^T r\left(\theta\right)\right)\right) \mathrm{d}\theta. \tag{1}$$

Given that $X \in \mathbb{N}_0^{n_\mathrm{X}}$, this is all we can do. However, if we make the assumption that the elements of X are "large" numbers so that $\left|\left(X\left(t + \Delta t\right) - X\left(t\right)\right)_j\right| \ll \left|X\left(t\right)_j\right|$, we can pretend that $X \in \mathbb{R}_0^{n_\mathrm{X}}$. If we also assume that the integrand is a continuous function of time, we can use the mean value theorem and rewrite the integral equations as differential equations,

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \dot{X}_\mathrm{i} - \dot{X}_\mathrm{e} + N \cdot \left(\nu^T r\right). \tag{2}$$

With similar assumptions, we can describe the total population as

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \dot{N}_\mathrm{i} - \dot{N}_\mathrm{e} + \dot{N}_\mathrm{g}. \tag{3}$$

By expressing X as $X = N \cdot \check{X}$, we find

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(N \cdot \check{X}\right) = \check{X}\frac{\mathrm{d}N}{\mathrm{d}t} + N\frac{\mathrm{d}\check{X}}{\mathrm{d}t}.$$

Likewise, we have

$$\dot{X}_\mathrm{i} \triangleq \dot{N}_\mathrm{i} \cdot \check{X}_\mathrm{i}$$
$$\dot{X}_\mathrm{e} \triangleq \dot{N}_\mathrm{e} \cdot \check{X}_\mathrm{e},$$

[14] $\mathbb{N}_0$ is the set of natural numbers including zero, i.e., $\{0, 1, 2, \ldots\}$. $\mathbb{Z}$ is the set of integers, i.e., $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$.

[15] $\mathbb{Q}$ is the set of rational numbers, which are *quotients* of integers.

where — assuming well mixed compartments — $\check{X}_\mathrm{e} \equiv \check{X}$. Thus, we find

$$\begin{aligned}
\check{X}\frac{\mathrm{d}N}{\mathrm{d}t} + N\frac{\mathrm{d}\check{X}}{\mathrm{d}t} &= \dot{X}_\mathrm{i} - \dot{X}_\mathrm{e} + N \cdot \left(\nu^\mathsf{T} r\right) \\
&\Downarrow \\
\check{X}\left(\dot{N}_\mathrm{i} - \dot{N}_\mathrm{e} + \dot{N}_\mathrm{g}\right) + N\frac{\mathrm{d}\check{X}}{\mathrm{d}t} &= \dot{N}_\mathrm{i} \cdot \check{X}_\mathrm{i} - \dot{N}_\mathrm{e} \cdot \check{X} + N \cdot \left(\nu^\mathsf{T} r\right) \\
&\Downarrow \\
N\frac{\mathrm{d}\check{X}}{\mathrm{d}t} &= \dot{N}_\mathrm{i}\left(\check{X}_\mathrm{i} - \check{X}\right) + N \cdot \left(\nu^\mathsf{T} r\right) - \dot{N}_\mathrm{g} \cdot \check{X} \\
&\Downarrow \\
\frac{\mathrm{d}\check{X}}{\mathrm{d}t} &= \frac{\dot{N}_\mathrm{i}}{N}\left(\check{X}_\mathrm{i} - \check{X}\right) + \nu^\mathsf{T} r - \frac{\dot{N}_\mathrm{g}}{N}\check{X}. \qquad (4)
\end{aligned}$$

Observe that if $N = \sum_j X_j$, then balance equations for every X_j in addition to a balance equation for N contains redundant information, with resulting possibilities of solving the differential equations for X and N leading to numerical inconsistencies. Normally, one might want to keep the balance equation for N, and remove one superfluous compartment balance X_s from vector X to a new $X \in \mathbb{N}_0^{n_\mathrm{X}-1}$, so that $X_\mathrm{s} = N - \sum_j X_j$. This way, numerical inconsistencies are avoided.

Often in epidemiology models, the total population is assumed to be constant, i.e., N = constant. Normally this is achieved by assuming $\dot{N}_\mathrm{i} \equiv \dot{N}_\mathrm{e} \equiv 0$, and in addition choosing either:

1. Dead individuals are included in the population, and births are neglected, or
2. Dead individuals are removed from the population, and the death rate equals the birth rate.

In both cases (1) and (2), this is equivalent to $\dot{N}_\mathrm{g} \equiv 0$. With these assumptions of N = constant, the model simplifies to

$$\frac{\mathrm{d}\check{X}}{\mathrm{d}t} = \nu^\mathsf{T} r.$$

The equivalent, non-scaled model is

$$\frac{\mathrm{d}X}{\mathrm{d}t} = N \cdot \left(\nu^\mathsf{T} r\right).$$

Whether we formulate the model in number of individuals, X, or fraction of individuals $\check{X}$, is not very important. However, it is often more convenient to consider fractional numbers because the results then generalize more easily to different populations.

2.3 Some general considerations

At the outset, number balances operate with integer states X, and the normal framework is the integral formulation in Eq. 1, where influents and effluents $\dot{X}_j$ as well as reaction events $\nu^\mathsf{T} r$ are zero except at a countable number of time instances t_k, when they are

integers. This implies that the model variables are integers and change at time instances which can be indexed by integers (countable time instances). We could denote such models as integer models: integer in value, and integer indexed time intervals. A brief introduction to such models is given, e.g., in Section 2.7.3 where the classical Reed-Frost model is discussed. There is also a First Reaction Time approach. Both of these integer models can be stochastic.

An alternative approach is to pretend that the variables X and time t are continuous variables, which allows us to pose differential equation number balances as in Eq. 2. This may lead to a number of people which is a real number (computes use a binary representation which really implies rational numbers). If we instead of posing number of people as unknown, instead pose the model with per capita variables as in Eq. 4, the variable are really rational numbers — which commonly are approximated by real numbers. Allowing for continuous time and real number variables is the standard approach in classical physics, and such models can be interpreted as deterministic differential equations. This is the approach used in the Problem section in this group project.

For completion, it is also possible to assume stochastic variables in the model. In that case, the model must be kept in integral form. The implication of including stochastic uncertainty is useful to be aware of.

2.4 The classical continuous SIR description

As in the classical Kermack & McKendrick (1927) model, we divide a population of N individuals into 3 groups: individuals of type S are *susceptible* for infection, individuals of type I are *infected*, while individuals of type R have *recovered* from the infection.

We have the chain of "reactions" $\mathrm{S} \to \mathrm{I} \to \mathrm{R}$, which can be broken down into the two independent parallel events $\mathcal{E}_j$, interaction with infection $\mathcal{E}_\mathrm{i}$ with rate r_i, and recovery $\mathcal{E}_\mathrm{r}$ with rate r_r:

$$\begin{aligned} \mathcal{E}_\mathrm{i} &: \mathrm{S} \overset{\mathrm{I}}{\to}{}^{k_\mathrm{i}} \mathrm{I}, \quad r_\mathrm{i} \\ \mathcal{E}_\mathrm{r} &: \mathrm{I} \to^{k_\mathrm{r}} \mathrm{R}, \quad r_\mathrm{r} \end{aligned}$$

where $\mathrm{S} \overset{\mathrm{I}}{\to}{}^{k_\mathrm{i}} \mathrm{I}$ indicates that an infected individual *catalyses* the transformation of a susceptible S into a new infected individual I without "consuming" the original infecting individual, with a frequency factor k_i or with a mean interval $\tau_\mathrm{i} = 1/k_\mathrm{i}$.[16]

In Keeling & Rohani (2008), this idea of "catalysis" where the available type I influences the event together with type S, without being consumed, is indicated as in Fig. 2.

It is important to stress that $\tau_\mathrm{i} = 1/k_\mathrm{i}$ is *not* the length of time an infected person stays infected (which probably is more or less constant), but rather the mean interval between each time one infected individual causes a susceptible to become infected. In other words, if a freshly infected person is removed from society (e.g., locked up, or otherwise put away), then $\tau_\mathrm{i} \to \infty$, alternatively $k_\mathrm{i} \to 0$, but the same infected person will still recover (or die) in finite time.

[16] Observe that in *chemical reaction engineering*, the phrase *frequency factor* is normally used for first order reactions, only. The reason is that in chemical engineering, reaction rates are normally given as $r = k \cdot f(c_j)$ where c_j is the *concentration* of species j. Since concentration has a dimension (material per volume), it is only for first order reactions that k has unit of *frequency* (= 1/time). Here, however, the rates are given as a function of dimensionless per capita numbers (equivalent to material *fraction* in chemical engineering), and then k has the unit of frequency.

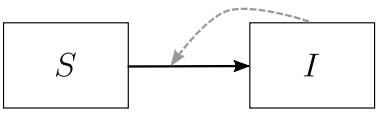

Figure 2: "Catalysis" of S by I, where presence of "catalyst" I influences the rate of transformation of S without without in itself transforming I. After Keeling & Rohani (2008), e.g., p. 16.

At the same time, $\mathrm{I} \to^{k_\mathrm{r}} \mathrm{R}$ indicates a simple recovery (i.e., no catalysis) from infected I to recovered R during a mean time of $\tau_\mathrm{r} = 1/k_\mathrm{r}$.

The stoichiometric reaction is

$$\underbrace{\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}}_{=\nu} \begin{pmatrix} \mathrm{S} \\ \mathrm{I} \\ \mathrm{R} \end{pmatrix} \leftarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

where ν is the *stoichiometric matrix*. It follows that

$$\begin{pmatrix} r_{\mathrm{S,g}} \\ r_{\mathrm{I,g}} \\ r_{\mathrm{R,g}} \end{pmatrix} = \nu^T \begin{pmatrix} r_\mathrm{i} \\ r_\mathrm{r} \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r_\mathrm{i} \\ r_\mathrm{r} \end{pmatrix} = \begin{pmatrix} -r_\mathrm{i} \\ r_\mathrm{i} - r_\mathrm{r} \\ r_\mathrm{r} \end{pmatrix}.$$

The infection rate r_i per capita is *catalyzed* by an infected individual, and takes place with a certain probability when a susceptible individual is in proximity of an infected individual, thus the probability of infection depends on the relative concentration of the two types,

$$r_\mathrm{i} \propto \frac{S}{N} \cdot \frac{I}{N} = \check{S} \cdot \check{I}.$$

This rate is in accordance with the *law of mass action* in chemical kinetics.[17] With time interval τ_i for this rate, we have

$$r_\mathrm{i} = \frac{1}{\tau_\mathrm{i}} \check{S}\check{I} = k_\mathrm{i} \check{S}\check{I}.$$

Similarly, reaction r_r is a gradual recovery rate which only depends on the concentration of the infected type, thus

$$r_\mathrm{r} = \frac{1}{\tau_\mathrm{r}} \check{I} = k_\mathrm{r} \check{I}.$$

For the SIR model presented here, ratio $\mathsf{R}_0 = \tau_\mathrm{r}/\tau_\mathrm{i} = k_\mathrm{i}/k_\mathrm{r}$ is known as the *basic reproduction number*.[18] The basic reproduction number thus is the ratio of the time constant of recovery and the time interval of infection. Essentially then, R_0 is the average number of people each individual infects.

Data for some epidemics are given in Table 1.

[17] Alternatively, the infection rate is sometimes posed as $N r_\mathrm{i} \propto S \cdot I$. In Martcheva (2015), this alternative formulation is incorrectly referred to as mass action based. With N constant, both infection rates $r_\mathrm{i} \propto \check{S} \cdot \check{I}$ and $r_\mathrm{i} \propto \frac{1}{N} S \cdot I$ give the same result, but when N varies, the results will be different.

[18] Observe: R_0 has nothing to do with the initial value R_0 of the recovered class.

Table 1: Basic reproduction number R_0, and recovery time $\tau_{\rm r} = 1/k_{\rm r}$. Columns 1, 2, and 5 are taken from https://en.wikipedia.org/wiki/Basic_reproduction_number.

Disease	R_0	$\tau_{\rm r}$ **[days]**	**Transmission**
Measles	12-18	7–14	Airborne
Polio	5-7		Fecal-oral route
Rubella	5-7	3	Airborne droplet
Mumps	4-7	5	Airborne droplet
Pertussis	5.5	60–90	Airborne droplet
Smallpox	3.5-6	30	Airborne droplet
HIV/AIDS	2-5		Body fluids
SARS	2-5		Airborne droplet
Diphtheria	1.7-4.3		Saliva
COVID-19	1.4-3.9	20–35	Airborne droplet
Influenza (1918)	1.4-2.8		Airborne droplet
Ebola (2014)	1.5-2.5		Body fluids
Influenza (2009)	1.4-1.6		Airborne droplet
Influenza(seasonal)	0.9-2.1	7	Airborne droplet
MERS	0.3-0.8		Airborne droplet

In Sanche et al. (2020), R_0 is estimated to be in the range $\mathsf{R}_0 \in [4.7, 6.6]$ as of late February 2020. As of mid March 2020, the European Centre for Disease Prevention and Control reports that early estimates of R_0 from Italy indicates R_0 to be in the interval $\mathsf{R}_0 \in [2.76, 3.25]$, with a reduction in R_0 shortly after the introduction of mitigation measures. A recent meta study reports the mean of R_0 at 3.28, with a median of 2.79, see [19]. It is probable that R_0 depends on the density of a population, and on the extent of commuting, etc. Furthermore, it is not clear to what degree these data include any level of mitigation measures.

In late March 2020, the Norwegian Public Health Institute[20] suggested $\mathsf{R}_0 = 2.4$ for light mitigation measures (quarantining people with positive COVID-19 diagnosis), and proposed $\mathsf{R}_0 = 1.3$ for the current measures (extended degree of "home office", closed schools, social distancing, etc.), and hoped for a future of $\mathsf{R}_0 = 0.9$ with even stronger measures (more testing, stricter enforcement of social distancing?).

2.5 Extension: the SEIR description

Now, we divide the population of N individuals into 4 groups: individuals of type S are *susceptible* for infection, individuals of type E are *exposed* to infection, but the infection is latent, individuals of type I are *infected*, while individuals of type R have *recovered* from the infection and are immune. Let S be the number of susceptible individuals, E be the number of exposed individuals, I be the number of infected individuals, and R be the number of recovered individuals. Neglecting birth and death, and neglecting influence and effluence of individuals, we have $N = S + E + I + R$. We have the chain of events $\mathrm{S} \to \mathrm{E} \to \mathrm{I} \to \mathrm{R}$, which can be broken down into the three parallel independent events

[19] https://www.ecdc.europa.eu/en/2019-ncov-background-disease

[20] Reported in newspaper Dagbladet, March 25, 2020.

$\mathcal{E}_\mathrm{i}$, $\mathcal{E}_\mathrm{e}$, and $\mathcal{E}_\mathrm{r}$ with rates r_i, r_e, and r_r, respectively:

$$\begin{aligned}
\mathcal{E}_\mathrm{i} &: \mathrm{S} \overset{\mathrm{I}}{\to}{}^{k_\mathrm{i}} \mathrm{E}, \qquad & r_\mathrm{i} &= \frac{1}{\tau_\mathrm{i}} \check{I}\check{S} = k_\mathrm{i}\check{I}\check{S} \\
\mathcal{E}_\mathrm{e} &: \mathrm{E} \to^{k_\mathrm{e}} \mathrm{I}, \qquad & r_\mathrm{e} &= \frac{1}{\tau_\mathrm{e}} \check{E} = k_\mathrm{e}\check{E} \\
\mathcal{E}_\mathrm{r} &: \mathrm{I} \to^{k_\mathrm{r}} \mathrm{R}, \qquad & r_\mathrm{r} &= \frac{1}{\tau_\mathrm{r}} \check{I} = k_\mathrm{r}\check{I}.
\end{aligned}$$

The stoichiometric reaction

$$\underbrace{\begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}}_{=\nu} \begin{pmatrix} \mathrm{S} \\ \mathrm{E} \\ \mathrm{I} \\ \mathrm{R} \end{pmatrix} \leftarrow \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

where ν is the *stoichiometric matrix*. It follows that

$$\begin{pmatrix} r_{\mathrm{S,g}} \\ r_{\mathrm{E,g}} \\ r_{\mathrm{I,g}} \\ r_{\mathrm{R,g}} \end{pmatrix} = \nu^T \begin{pmatrix} r_\mathrm{i} \\ r_\mathrm{e} \\ r_\mathrm{r} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_\mathrm{i} \\ r_\mathrm{e} \\ r_\mathrm{r} \end{pmatrix} = \begin{pmatrix} -r_\mathrm{i} \\ r_\mathrm{i} - r_\mathrm{e} \\ r_\mathrm{e} - r_\mathrm{r} \\ r_\mathrm{r} \end{pmatrix}.$$

2.6 The classical discrete SIR description

The classical *Reed-Frost* model, Schwabe et al. (1977), was based on the SIR compartmentalization, was discrete in time, and assumed a stochastic infection transmission with a fixed probability. To some degree, this model is related to agent-based models, where one could envision tracking of individuals within a compartment.

Each of the I_t infected individuals at time period t is in contact with every susceptible S_t with a probability p of infecting the other individual. An infected individual then recovers before time period $t+1$. This implies that the time interval between each time period in the discrete model must be in the order of the recovery time τ_r. It furthermore means that those who are infected in time period $t+1$ became infected in time period t.

The model is based on the following assumptions:

- The infection is spread directly from infected individuals due to "adequate contact".
- A susceptible individual who contracts an infection in a given time period (t) will be infected in the next time period ($t+1$), and will recover and become immune in every subsequent time period ($t+2, t+3, \ldots$).
- Each susceptible individual has a fixed probability p of infection ("adequate contact") from any infected individual in the population within the time period. The probability of infection p is the same for every individual.
- The population size N is constant, and there is no contact with individuals outside of the population.
- All conditions are constant during the epidemic.

The following parameters and conditions are needed at start-up:

1. Size of population, N.
2. Number of infected, I_0 — often set to $I_0 = 1$.
3. Number of already immune individuals, R_0. The initial number of susceptibles is then $S_0 = N - I_0 - R_0$.
4. The probability p of infection ("adequate contact").

By comparison with the deterministic SIR model, we have

$$\mathsf{R}_0 = S_0 \cdot p \Rightarrow p = \frac{\mathsf{R}_0}{S_0}.$$

2.7 Dynamics of epidemics

2.7.1 Dynamic SIR model

Combining these expressions in the number balance, we find the following model for the number of susceptibles, infectious, and recovered (known as the SIR model):

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \dot{S}_\mathrm{i} - \dot{S}_\mathrm{e} - \frac{k_\mathrm{i}}{N} IS \tag{5}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \dot{I}_\mathrm{i} - \dot{I}_\mathrm{e} + \frac{k_\mathrm{i}}{N} IS - k_\mathrm{r} I \tag{6}$$

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \dot{R}_\mathrm{i} - \dot{R}_\mathrm{e} + k_\mathrm{r} I. \tag{7}$$

Neglecting immigration ($\dot{S}_\mathrm{i} = \dot{I}_\mathrm{i} = \dot{R}_\mathrm{i} \equiv 0$) and emmigration ($\dot{S}_\mathrm{e} = \dot{I}_\mathrm{e} = \dot{R}_\mathrm{e} \equiv 0$), and with constant total number of people N, $N = S + I + R$, we can simplify the model, and also write it in per capita form:

$$\frac{\mathrm{d}\check{S}}{\mathrm{d}t} = -k_\mathrm{i} \check{I} \check{S}$$

$$\frac{\mathrm{d}\check{I}}{\mathrm{d}t} = k_\mathrm{i} \check{I} \check{S} - k_\mathrm{r} \check{I}$$

$$\frac{\mathrm{d}\check{R}}{\mathrm{d}t} = k_\mathrm{r} \check{I},$$

or slightly rewritten:

$$\frac{\mathrm{d}\check{S}}{\mathrm{d}t} = -k_\mathrm{i} \check{I} \check{S} \tag{8}$$

$$\frac{\mathrm{d}\check{I}}{\mathrm{d}t} = \left(k_\mathrm{i} \check{S} - k_\mathrm{r}\right) \check{I} \tag{9}$$

$$\frac{\mathrm{d}\check{R}}{\mathrm{d}t} = k_\mathrm{r} \check{I}. \tag{10}$$

Under our assumptions with $N = S + I + R =$ constant or $1 = \check{S} + \check{I} + \check{R}$, and we really only have 2 independent differential equations. Numerically, it is thus best to drop the differential equation for $\check{R}$ and instead compute $\check{R}$ as $\check{R} = 1 - \check{S} - \check{I}$.

2.7.2 Dynamic SEIR model

Combining the relevant expressions for the SEIR reactions in the number balance, we find the following SEIR model:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \dot{S}_{\mathrm{i}} - \dot{S}_{\mathrm{e}} - \frac{k_{\mathrm{i}}}{N} IS \tag{11}$$

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \dot{E}_{\mathrm{i}} - \dot{E}_{\mathrm{e}} + \frac{k_{\mathrm{i}}}{N} IS - k_{\mathrm{e}} E \tag{12}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \dot{I}_{\mathrm{i}} - \dot{I}_{\mathrm{e}} + k_{\mathrm{e}} E - k_{\mathrm{r}} I \tag{13}$$

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \dot{R}_{\mathrm{i}} - \dot{R}_{\mathrm{e}} + k_{\mathrm{r}} I, \tag{14}$$

or without immigration and emigration, with constant number of people N, and in per capita form:

$$\begin{aligned}
\frac{\mathrm{d}\check{S}}{\mathrm{d}t} &= -k_{\mathrm{i}} \check{I} \check{S} \\
\frac{\mathrm{d}\check{E}}{\mathrm{d}t} &= k_{\mathrm{i}} \check{I} \check{S} - k_{\mathrm{e}} \check{E} \\
\frac{\mathrm{d}\check{I}}{\mathrm{d}t} &= k_{\mathrm{e}} \check{E} - k_{\mathrm{r}} \check{I} \\
\frac{\mathrm{d}\check{R}}{\mathrm{d}t} &= k_{\mathrm{r}} \check{I}.
\end{aligned}$$

Under our assumptions with $1 = \check{S} + \check{E} + \check{I} + \check{R} =$ constant, we really only have 3 independent differential equations, and should compute the fourth value from the number constancy.

2.7.3 Dynamic discrete SIR model

According to the Reed-Frost model,

$$I_{t+1} = \mathbf{B}\left(S_t, 1 - (1-p)^{I_t}\right) \tag{15}$$

$$S_{t+1} = S_t - I_{t+1} \tag{16}$$

$$R_{t+1} = N - S_{t+1} - I_{t+1}. \tag{17}$$

Here, $\mathbf{B}(n, x)$ is a random number in the set $\{0, \ldots, n\}$ drawn from the *Binomial distribution* at specified probability x.[21]

The model is developed as follows. The probability of *non*-adequate contact between a susceptible S and a single infected I is $1 - p$. Thus, the probability of *non*-adequate contact between a susceptible S and *every* infected individuals in time period t is $(1-p)^{I_t}$. The probability of *adequate* contact between susceptible S and every infected individual in time period t is thus $1 - (1-p)^{I_t}$.

In a stochastic set-up, the chance that *one* susceptible S contracts an infection in time period t is thus $\mathbf{B}_1\left(1, 1 - (1-p)^{I_t}\right)$, where $\mathbf{B}_1(1, x)$ is the *Bernoulli* distribution.[22] The

[21] Here, $\mathbf{B}(\cdot)$ is short for $\mathbf{Binomial}(\cdot)$.
[22] Here, $\mathbf{B}_1(\cdot)$ is short for $\mathbf{Bernoulli}(\cdot)$.

Bernoulli distribution $\mathbf{B}_1(1, x)$ has outcome either 1 (true) or 0 (false): if we draw from this distribution infinitely many times, the fraction with outcome 1 is x.

As S_t susceptibles move around in the population of I_t infected, each with a probability x of adequate contact with some infected, the total number of infected individuals becomes $\mathbf{B}\left(S_t, 1-(1-p)^{I_t}\right)$. Because infection only lasts one time period, Eq. 15 follows. Equations 16–17 follow immediately.

It can be shown that the *expected* infection population expression in Eq. 15

$$I_{t+1} = S_t\left(1-(1-p)^{I_t}\right). \tag{18}$$

From here, it follows that the *expected* susceptible population is

$$S_{t+1} = S_t - S_t\left(1-(1-p)^{I_t}\right) = S_t(1-p)^{I_t}, \tag{19}$$

while Eq. 17 is still valid for the recovered individuals R.

The discrete time/stochastic SIR model can also be expressed in fractional terms, and then becomes

$$\begin{aligned}
\check{I}_{t+1} &= \frac{1}{N}\mathbf{B}\left(\check{S}_t N, 1-(1-p)^{\check{I}_t N}\right)\\
\check{S}_{t+1} &= \check{S}_t - \check{I}_{t+1}\\
\check{R}_{t+1} &= 1 - \check{S}_{t+1} - \check{I}_{t+1}.
\end{aligned}$$

Here, it is important that $\check{S}_t N$ and $\check{I}_t N$ are integers, so it is safer to pose the model using number of individuals instead of fractions of individuals.

The approach with a given constant (or time varying) probability of infection is similar to what is used in agent-based models (material description) where each individual or group of individuals is tracked. In more general agent-based models, the time interval does not have to be assume recovery within the time interval.

The dynamic discrete SIR model is *illustrated* in a Measles case study, but is not part of the project work.

2.7.4 Continuous time stochastic SIR models

In addition to the discrete time stochastic models based on Bernoulli distribution of the infection at given probability, it is possible to consider "continuous" time stochastic models.

Continuous time stochastic models are based on a Poisson distribution of events related to independent rates of change. These models can either be posed as stochastic differential equations (SDEs) with continuous variation in the population numbers (i.e., $S, I, R \in \mathbb{R}$) where the Poisson distribution typically is approximated by a normal distribution, or instead $S, I, R \in \mathbb{N}_0$ and the time *between* events can be shown to follow a Exponential distribution.

These continuous time stochastic SIR models are *illustrated* in a Measles case study, but are not part of the project work.

Table 2: Daily number of influenza infected at boarding school with 763 boys in Northern England, January-February 1978. Taken from Martcheva (2015).

Day	# infected	Day	# infected
3	25	9	192
4	75	10	126
5	227	11	71
6	296	12	28
7	258	13	11
8	236	14	7

2.8 Measles case study

2.8.1 Overview of case study

At a boarding school in North England in January–February 1978, measles infection was observed among the 763 pupils, with 25 infections on assumed day 3 of the epidemic. The evolution of observed infections was as recorded in Table 2.

In Martcheva (2015), a SIR model is used to model the infection, assuming that $N = 763$ is constant, and that all pupils are locked-up in the school.

2.8.2 Continuous time: deterministic dynamic model

The model in Eqs. 5–7 with neither immigration nor emmigration, and with constant number of pupils, can be implemented and run in Julia as follows. The following Julia packages are used.

```
using Plots; pyplot()
using LaTeXStrings
using DifferentialEquations
```

The following constants have been defined:

```
# Linewidths and styles
LW1 = 2.5
LW2 = 1.5
LS1 = :solid
LS2 = :dot
LS3 = :dash
LA1 = 1
LA2 = 0.7
LA3 = 0.4
LA4 = 0.2
MS1 = 7
;
```

The SIR model is implemented in the following function:

```
# Function for SIR model
function sir_expanded!(dx,x,p,t,N)
    S,I,R = x
```

```
    k_i,k_r = p
    #
    dS = -k_i*I*S/N
    dI = k_i*I*S/N - k_r*I
    dR = k_r*I
    dx .= [dS,dI,dR]
end
#
sir!(dx,x,p,t) = sir_expanded!(dx,x,p,t,N)
```

Martcheva (2015), p. 127 proposes original model parameters[23]

$$k_{\mathrm{i}} = \beta N = 0.0025 \cdot 763 = 1.9075$$
$$k_{\mathrm{r}} = \alpha = 0.3.$$

Data and set-up of the differential equations problem, and solution of the model:

```
# Infection data
N = 763
tm = 3:14
Id = [25, 75, 227, 296, 258, 236, 192, 126, 71, 28, 11, 7]
# Initial values of model
I3 = Float64(Id[1])
S3 = N-I3
R3 = 0.
x3 = [S3,I3,R3]
# Model parameters
k_i = N*0.0025
k_r = 0.3
RR0 = round(k_i/k_r,digits=1)
p = [k_i, k_r]
# Time span for simulation
tspan=(3.,14.)
# Problem definition and solution
prob = ODEProblem(sir!,x3,tspan,p)
sol = solve(prob)
```

Plotting of results:

```
plot(sol,lw=LW1,lc=sirCOL,label=["\$S,␣\\mathsf{R}_0=$RR0\$" "\$I,␣\\
    mathsf{R}_0=$RR0\$" "\$R,␣\\mathsf{R}_0=$RR0\$"])
scatter!(tm, Id, st=:scatter,ms=MS1,mc=sirCOL[2],label=L"I^\mathrm{d}")
plot!(xlabel="time␣[day]",ylabel="#␣boys",title="Boarding␣school␣measles
    ␣infection")
```

The result of the simulation with original parameters vs. the registered infection data are as in Fig. 3.

Observe that the effective basic reproduction number R_0 with original parameters is found to be $\mathsf{R}_0 \approx 6.4$, which is lower than the range proposed in Table 1. The reason is that $\mathsf{R}_0 \approx 6.4$ assumes some mitigation, e.g., restricting social contact with infected pupils.

[23] Martcheva (2015) denotes the model parameters by α and β.

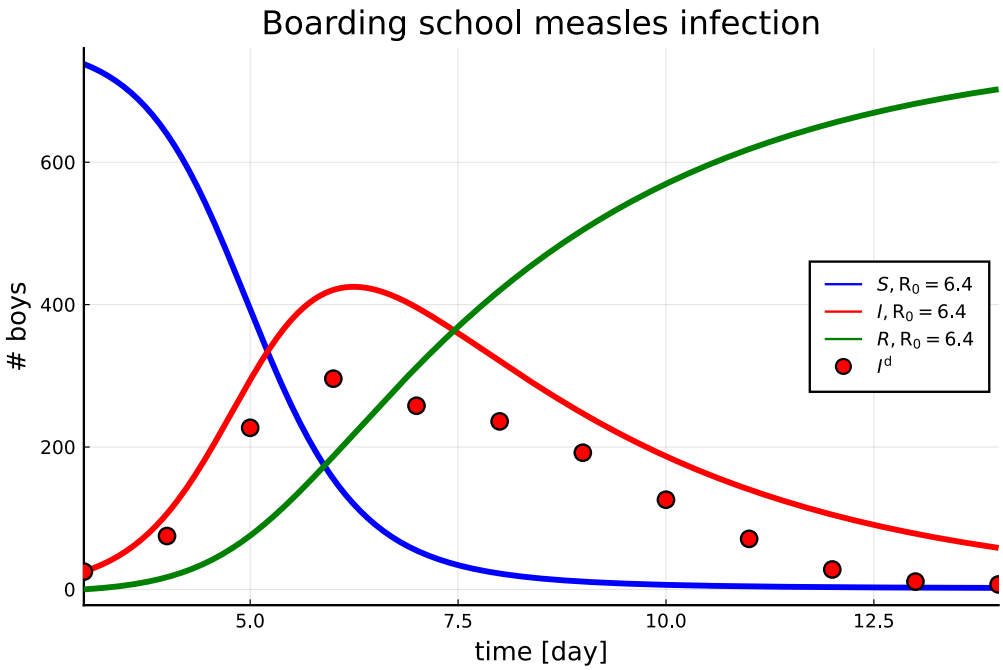

Figure 3: Comparing SIR model with original parameters vs. registered infection data for boarding school in North England in January–February 1978.

2.8.3 Deterministic dynamic model: model fitting

It is of interest to find better model parameters. Model fitting techniques go beyond the scope of course FM1015. Simulation of the model with the fitted parameters is displayed in Fig. 5 with dark/thick lines.

It is also of interest to have an idea of parameter uncertainty in this fitted model. Figure. 4 shows the probability distribution in k_i vs. k_r as found using the Markov Chain Monte Carlos method.

Point estimators (mean values) of the parameters are found to be:

$$k_\mathrm{i} \approx 1.817$$
$$k_\mathrm{r} \approx 0.4618.$$

Other choices of point estimators are also possible, e.g., the median values, the values that maximize the parameter probability density function (in Fig. 4), etc. — these values may differ.

It is also of interest to have an idea of how the parameter uncertainty translates into prediction uncertainty; known as *data retrodiction*[24]. Results are displayed in Fig. 5 in pale/thin lines, together with the simulation when using the point estimates in dark/thick lines.

From the data retrodiction, we see that there is relatively little uncertainty in the

[24] *retrodiction*, from Latin: *retro* = backward, back, behind + *dicere* = to say, tell, speak. In classical Latin, the verb *dicere* had the "c" pronounced as "k", e.g., approximately as *diquere* would be pronounced in English. In the later "Church Latin" (medieval Latin), the pronounciation of "c" changed to a "ch" sound (as in modern Italian), e.g., approximately as *dichere* would be pronounced in English. *Prediction* is to say in advance, *Data retrodiction* is to say *after* the model has been fitted to data.

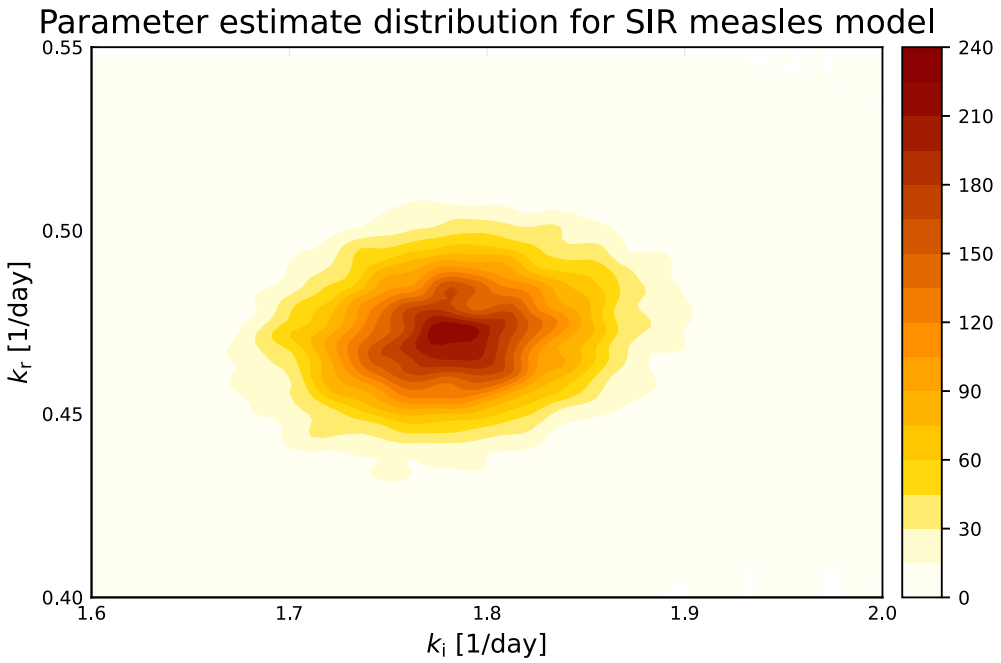

Figure 4: Distribution in Bayes estimates of parameters k_i and k_r in SIR model based on infection data for boarding school in North England in January–February 1978.

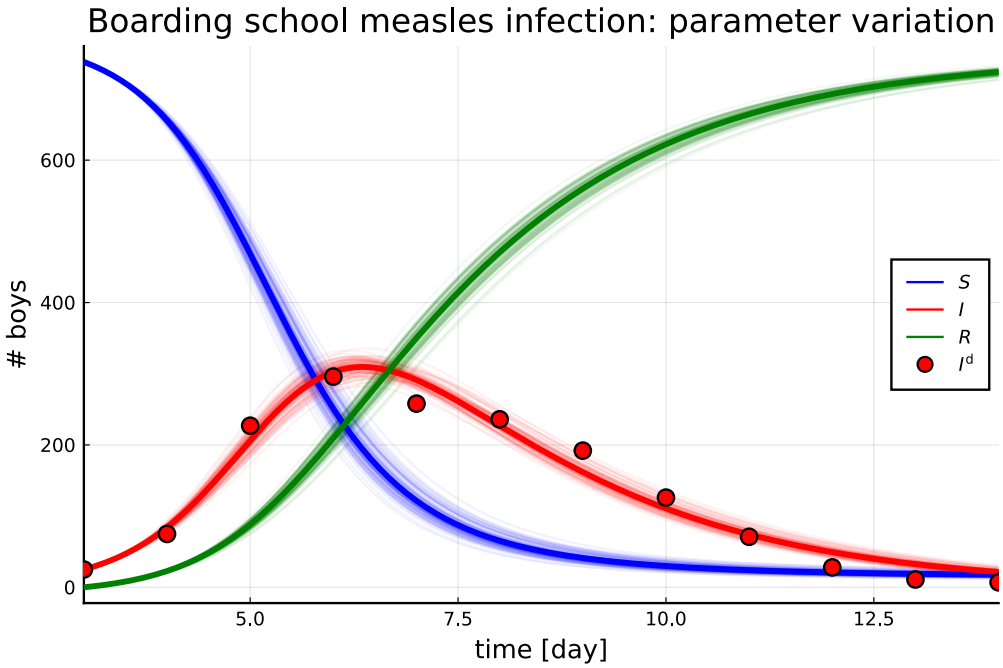

Figure 5: Comparing SIR model *data retrodiction* vs. registered infection data for boarding school in North England in January–February 1978. Dark/thick lines are based on the point estimates.

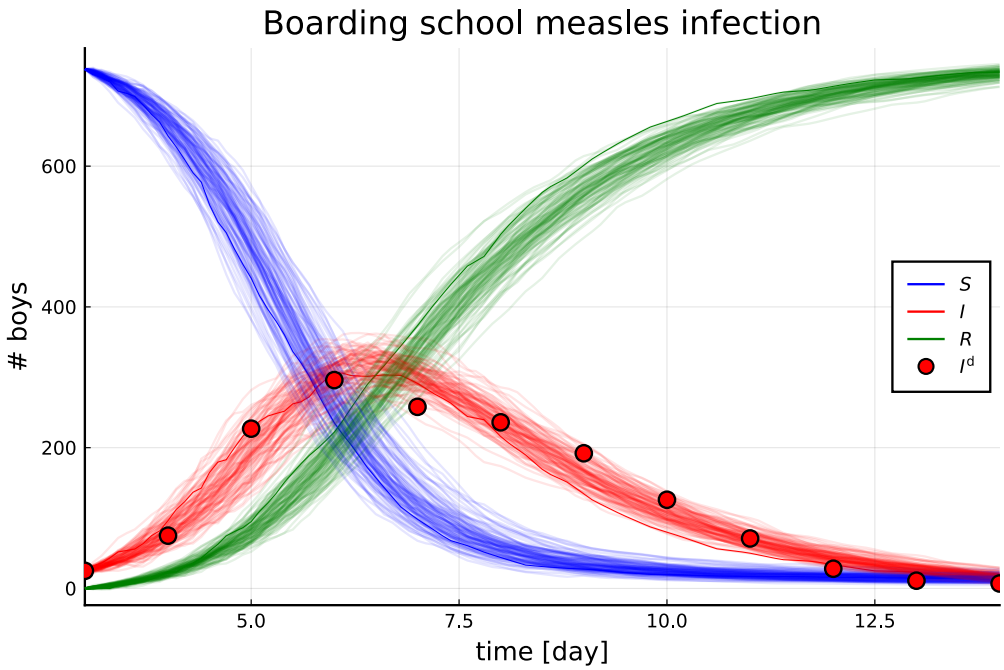

Figure 6: Stochastic realizations (trajectories) for an ensemble of 100 possible scenarios.

model even with varying parameters. This indicates that the model has quite good predictive properties.

2.8.4 Continuous time: stochastic differential equation model

Using basic data from the deterministic simulation, we expand the model to a set of stochastic differential equations. We can now formulate a model consisting of a drift term and a diffusion term. This goes beyond the content of course FM1015. The results are shown in Fig. 6.

The results are shown in Fig. 7.

2.8.5 Continuous time: stochastic first reaction event model

Instead of formulating the model as Stochastic Differential Equations, we can write integer difference equations with first reaction event description for changes. This goes beyond the content of course FM1015.

The results are shown in Fig. 8, and should be compared to the results in Fig. 6.

It is of interest to compute some statistics from the realizations. The results are shown in Fig. 9, and should be compared to the results in Fig. 7.

2.8.6 Discrete time: stochastic SIR model

Finally, we consider the discrete time Reed-Frost SIR model. This model is quite limited in the sense that we need to specify the period for which everyone has recovered, τ_{p}; observe that this period time τ_{p} is not necessarily the same as $\tau_{\mathrm{r}} = 1/k_{\mathrm{r}}$ in the continuous time SIR models. To get good fit with experimental data, it is often necessary to choose the discrete time interval τ_{p} to quite large, hence S, I , and R are known in rather few time instances. Here, we use probability of infection given by $p = \frac{\mathrm{R}_0}{S_3} \approx 5.285 \cdot 10^{-3}$ and

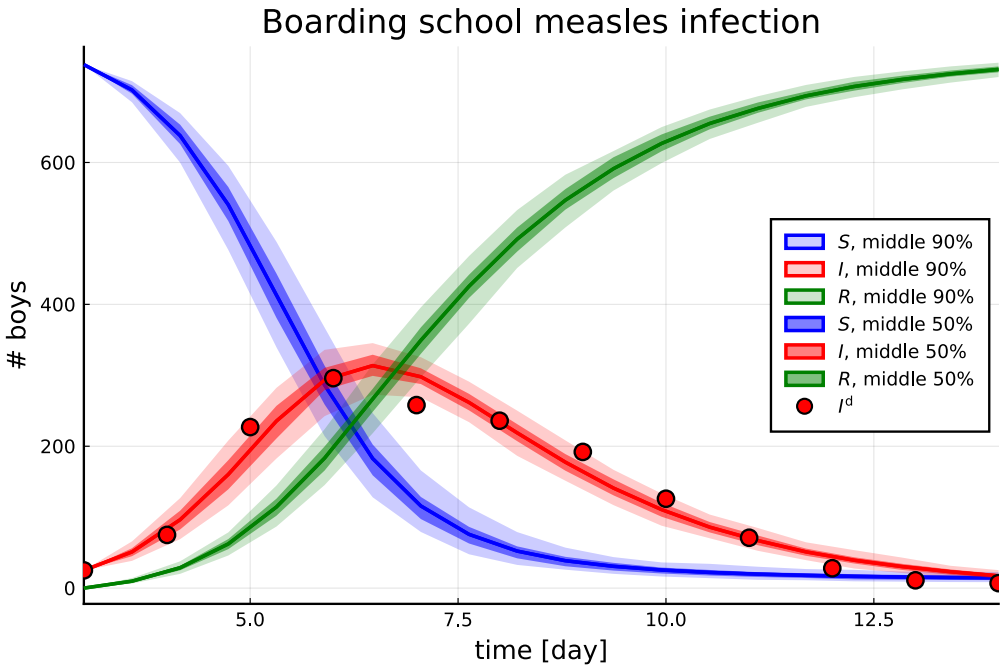

Figure 7: Statistics of realizations (trajectories) for an ensemble of 100 possible scenarios.

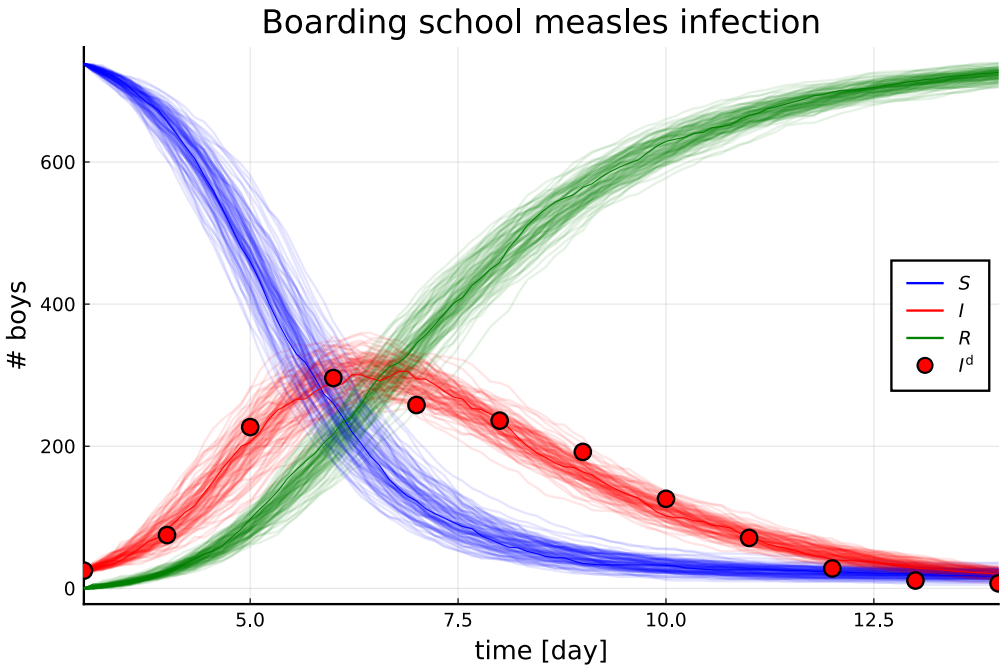

Figure 8: Stochastic realizations (trajectories) for an ensemble of 100 possible scenarios using Gillespie's First Reaction Method (resampled for data reduction).

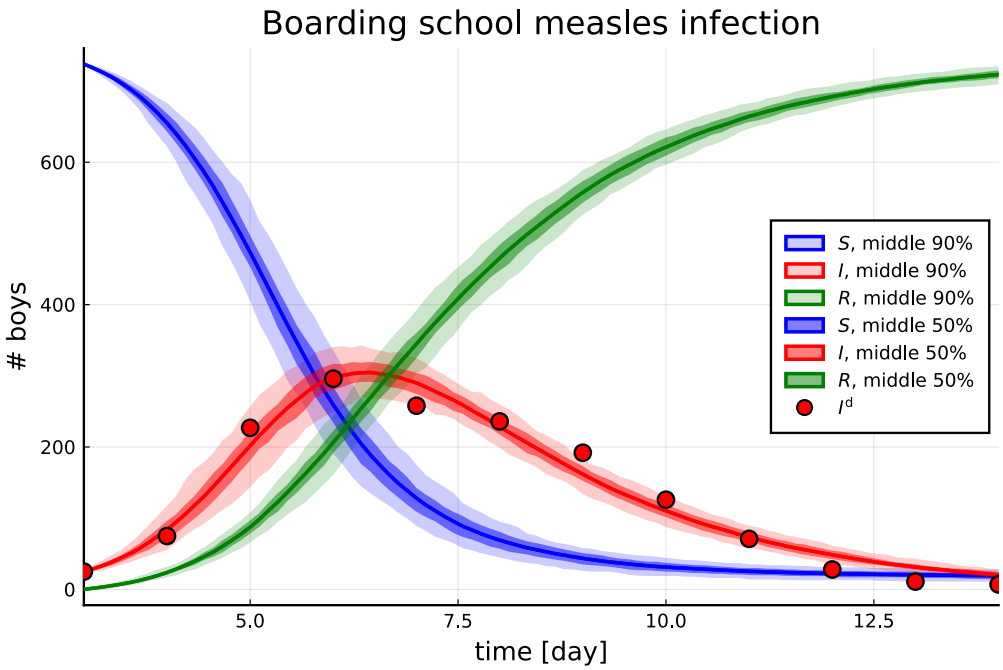

Figure 9: Statistics of realizations (trajectories) for an ensemble of 100 possible scenarios based on Gillespie's First Reaction Method.

set the recovery time to $\tau_{\mathrm{p}} = 1.8$ for an initial simulation, and then re-tune these two parameters to better fit the data points using least squares fit. The results of the refitted parameters are

$$\begin{aligned} \tau_{\mathrm{p}} &\approx 1.976 \\ p &\approx 6.53 \cdot 10^{-3}. \end{aligned}$$

The results are shown in Fig. 10, and should be compared to the results in Figs. 6 and 8.

It is of interest to compute some statistics from the realizations. The results in Fig. 11 should be compare with Figs. 7 and 9.

2.8.7 Continuous time: deterministic SEIR model

The model in Eqs. 11–14 with neither immigration nor emmigration, and with constant number of pupils, can also be used. Here, we will assume that I_3 is known, that $R_3 = 0$, that N is known, but that S_3 and E_3 are unknown, so we will need to estimate, say S_3.[25] We reuse the model parameters for k_{i} and k_{r} before model fitting for the SIR model, and first assume that $k_{\mathrm{e}} = 2$; Fig. 12.

We can also fit the parameters of the SEIR model, including the initial value of S_3.

[25]Typically, the initial value of the chosen state (S_3 or E_3) will be assumed to have some (truncated) normal distribution. Because S_3 and E_3 should really be non-negative integers, it is safer to estimate the initial value of S_3 than E_3, because S_3 is quite large with little chance of drawing a negative number, while E_3 is close to zero with a substantial chance of drawing a negative number.

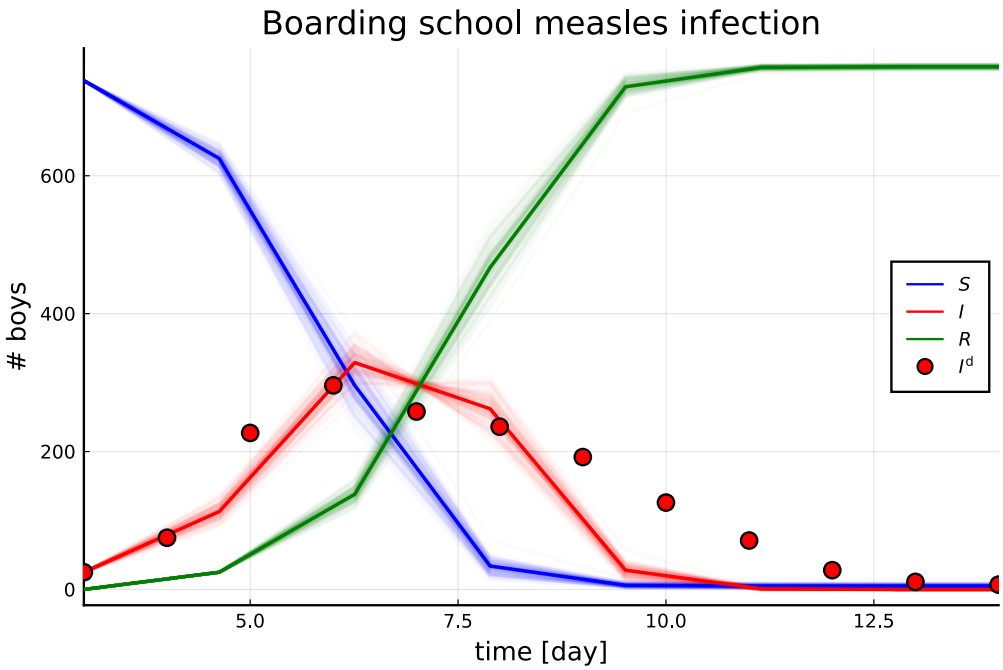

Figure 10: Deterministic simulation of the Reed-Frost SIR model (dark lines) and stochastic realizations (lighter colored trajectories) for an ensemble of 100 possible scenarios using the stochastic Reed-Frost SIR model.

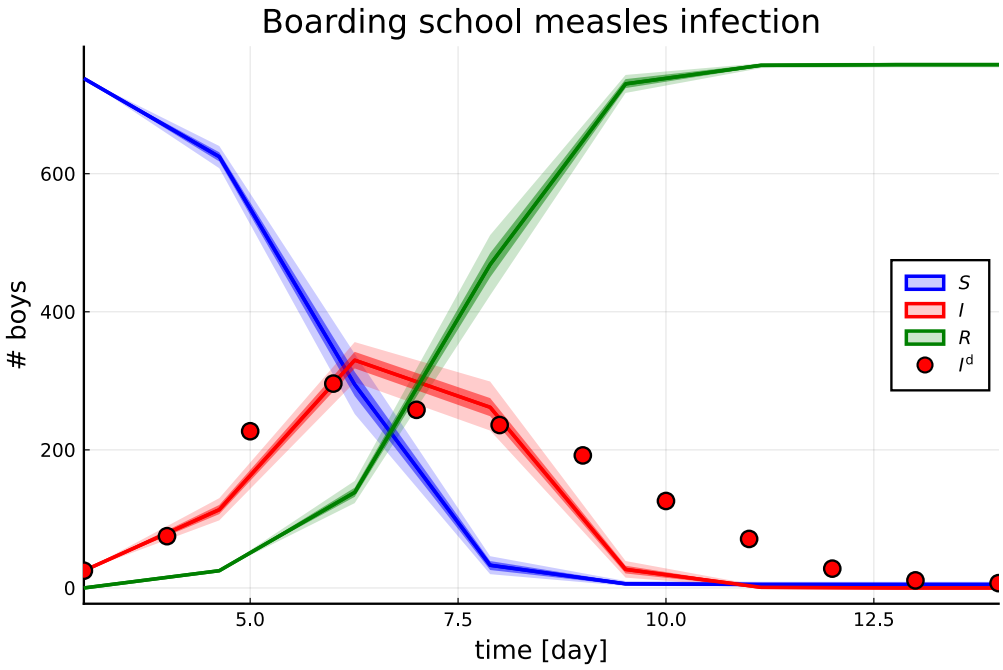

Figure 11: Statistics of realizations (trajectories) for an ensemble of 100 possible scenarios based on the stochastic discrete time Reed-Frost SIR model.

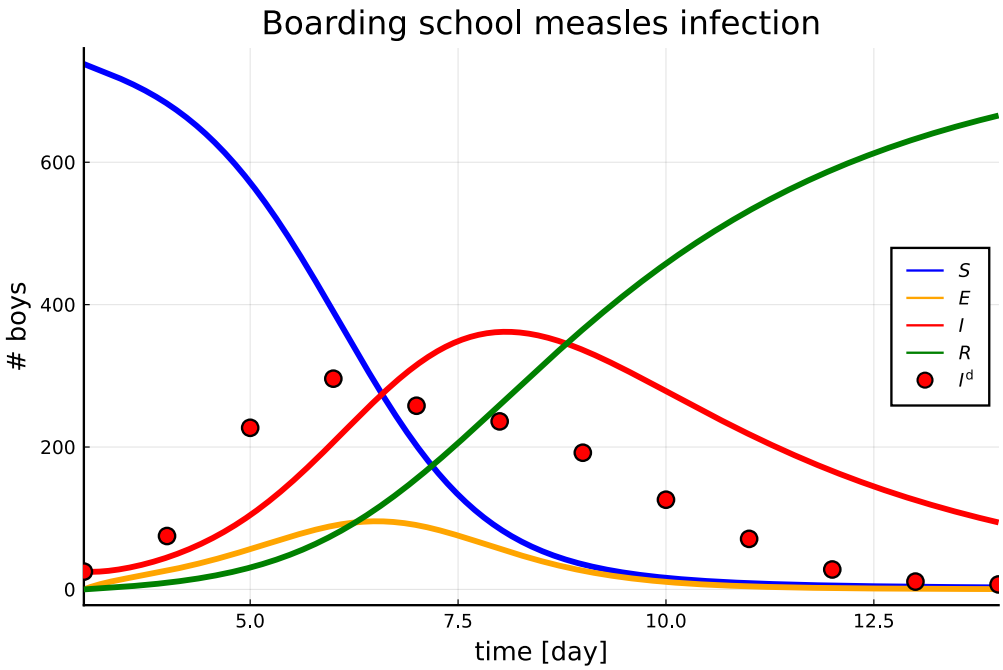

Figure 12: Comparing SEIR model with original parameters vs. registered infection data for boarding school in North England in January–February 1978.

The result are the following point estimates (mean value):

$$
\begin{aligned}
k_\mathrm{i} &\approx 2.50 \\
k_\mathrm{e} &\approx 2.826 \\
k_\mathrm{r} &\approx 0.477 \\
S_3 &\approx 720
\end{aligned}
$$

The distribution in parameters k_i vs. k_r is illustrated in Fig. 13.

The distribution in parameters k_i vs. k_e is illustrated in Fig. 14.

The distribution in parameters k_e vs. k_r is illustrated Fig. 15.

If we draw estimates of these parameters and initial value and redo simulations, the data retrodiction is as in Fig. 16. As we see, the uncertainties in the retrodiction of the SEIR model, Fig. 16, are far larger than the uncertainties in the retrodiction of the SIR model, Fig. 5. This may indicate that the measles infection is best modeled by a SIR model, but it must be remembered that we have few data points and the SEIR model has almost twice as many parameters as the SIR model[26].

[26]Estimated parameters for the SIR model: k_i, k_r, and standard deviation in output error. Estimated parameters for the SEIR model: k_i, k_e, k_r, initial value for S_3, and standard deviation in output error.

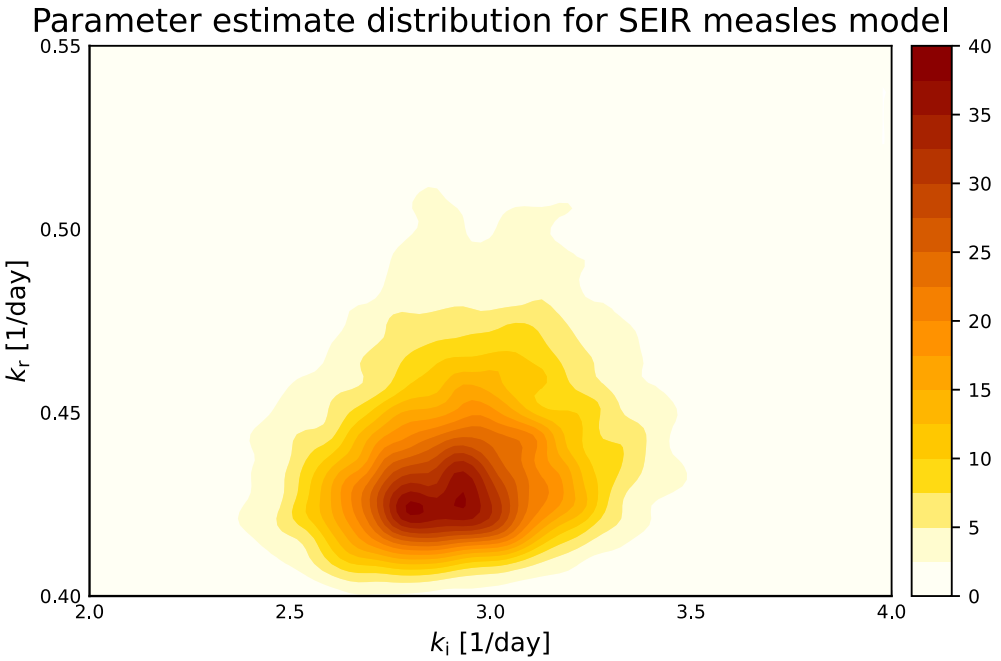

Figure 13: Distribution in Bayes estimates of parameters k_i and k_r in SIR model based on infection data for boarding school in North England in January–February 1978.

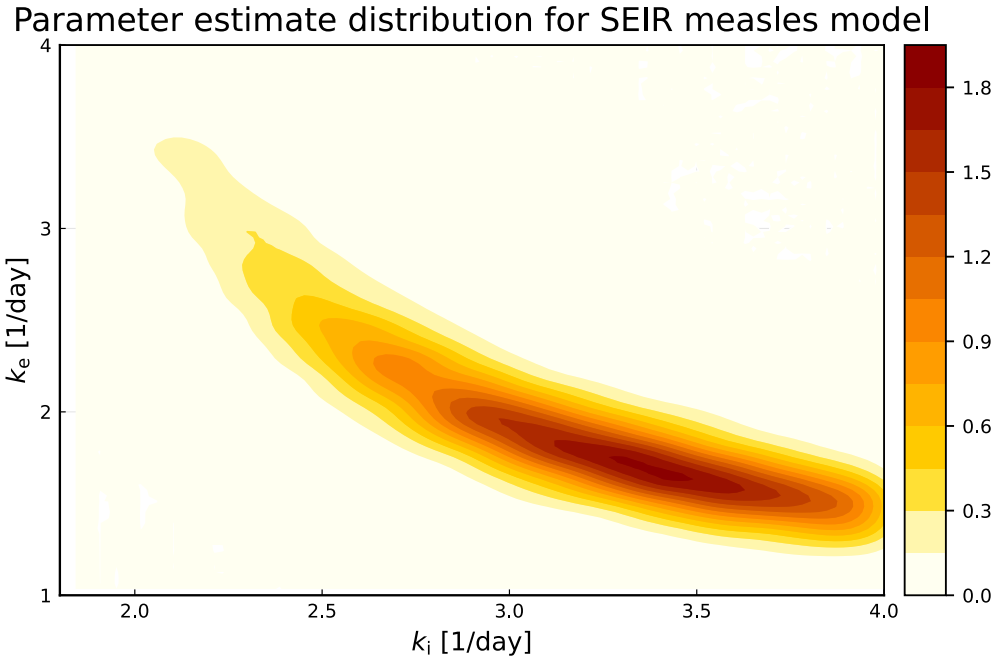

Figure 14: Distribution in Bayes estimates of parameters k_i and k_r in SIR model based on infection data for boarding school in North England in January–February 1978.

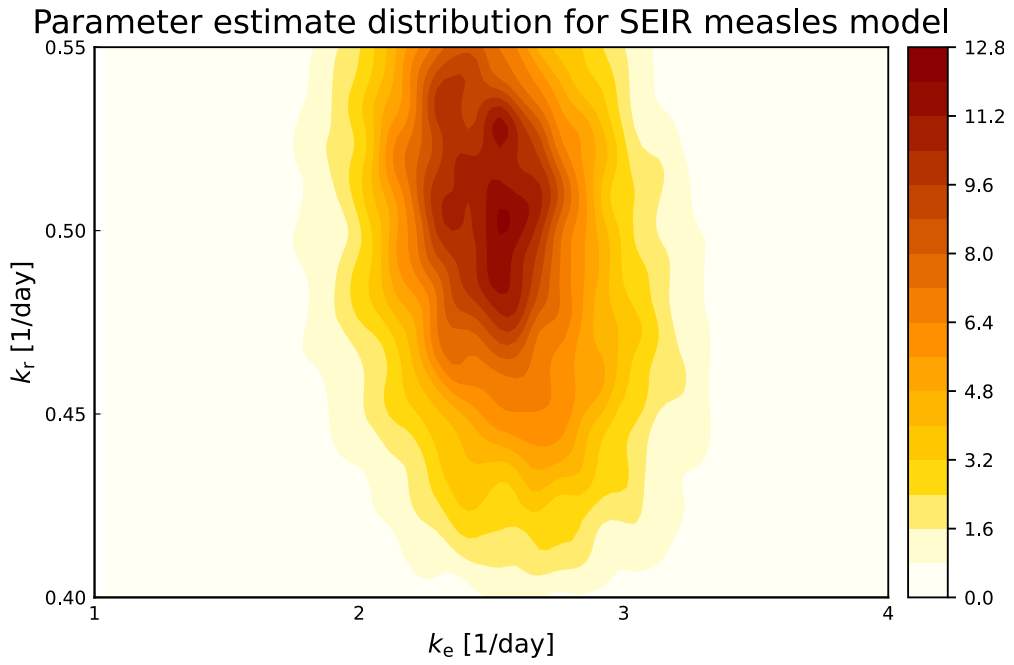

Figure 15: Distribution in Bayes estimates of parameters k_i and k_r in SIR model based on infection data for boarding school in North England in January–February 1978.

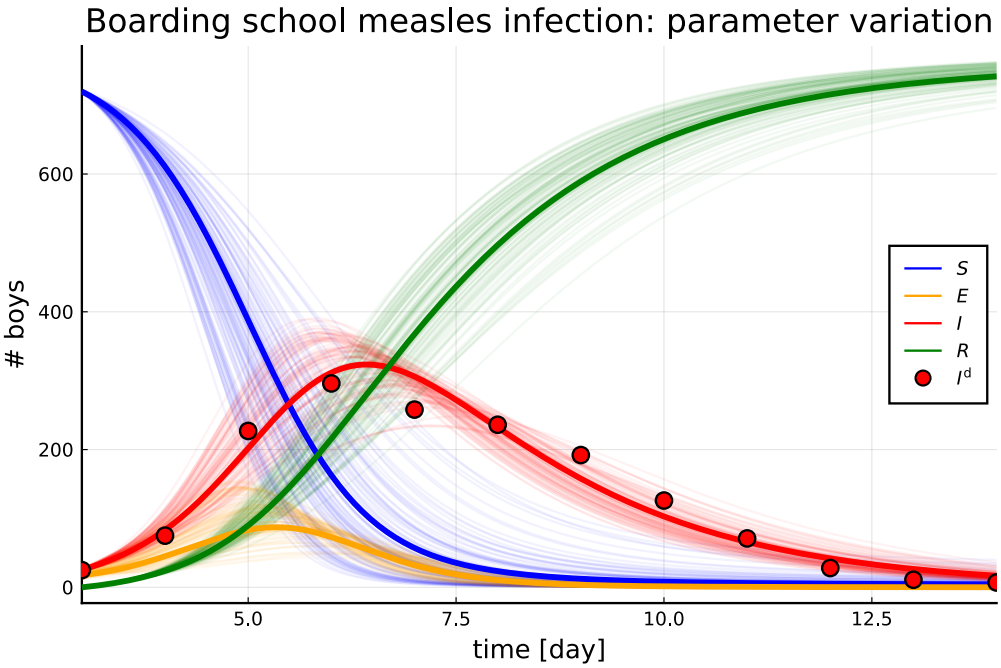

Figure 16: Comparing SEIR model data *retrodiction* vs. registered infection data for boarding school in North England in January–February 1978. Dark lines are based on the point estimates.

3 Analysis of epidemiology models

3.1 SIR: equilibrium solutions

3.1.1 The case of no birth/death

Analysis of epidemiology models is simplest to do for constant population N without immigration or emigration, and based on per capita models. For the SIR model in Eqs. 8–10, the equilibrium point(s) are found by setting derivatives to zero, leading to

$$\begin{aligned} -k_{\mathrm{i}}\check{I}\check{S} &= 0 \\ \left(k_{\mathrm{i}}\check{S} - k_{\mathrm{r}}\right)\check{I} &= 0 \\ \check{I} &= 0. \end{aligned}$$

From the first of these equations, we see that equilibrium requires either $\check{I} = 0$ or $\check{S} = 0$ or both, the second equation requires $\check{I} = 0$ or $\check{S} = \frac{k_{\mathrm{r}}}{k_{\mathrm{i}}} = \frac{\tau_{\mathrm{i}}}{\tau_{\mathrm{r}}} = \frac{1}{\mathsf{R}_0}$ or both, while the third unequivocally requires $\check{I} = 0$. Thus, $\check{I}$ has to be zero in equilibrium.

Because equilibrium requires that $\check{I} \equiv 0$, we see that equilibrium is satisfied for any values of $\check{S}$ and $\check{R}$ as long as $\check{S} + \check{R} = 1$. One possibility is that $\check{S} = 1$ and $\check{R} = 0$ (the *disease-free* case). Thus,

$$\begin{aligned} E_0: &\quad \left(\check{S}, \check{I}, \check{R}\right) = (1, 0, 0) \\ E_1: &\quad \left(\check{S}, \check{I}, \check{R}\right) = \left(\check{S}, 0, 1 - \check{S}\right) \end{aligned}$$

where E_0 is the disease free case.

3.1.2 The case of birth/death

In slightly more realistic models, it is common to introduce a birth rate $\dot{\beta}$ and a death rate $\dot{\delta}$. To preserve constant population, it is necessary that $\dot{\beta} \equiv \dot{\delta}$. People are often assumed to be born susceptible, with equal probability of death from each of the classes S, I, and R. The per capita SIR model then changes to

$$\frac{\mathrm{d}\check{S}}{\mathrm{d}t} = \dot{\beta} - k_{\mathrm{i}}\check{I}\check{S} - \dot{\beta}\check{S} \tag{20}$$

$$\frac{\mathrm{d}\check{I}}{\mathrm{d}t} = \left(k_{\mathrm{i}}\check{S} - k_{\mathrm{r}}\right)\check{I} - \dot{\beta}\check{I} \tag{21}$$

$$\frac{\mathrm{d}\check{R}}{\mathrm{d}t} = k_{\mathrm{r}}\check{I} - \dot{\beta}\check{R}. \tag{22}$$

Equilibrium now requires

$$\begin{aligned} 0 &= \check{S}\left(\dot{\beta} + k_{\mathrm{i}}\check{I}\right) - \dot{\beta} \\ 0 &= \check{I}\left(k_{\mathrm{i}}\check{S} - k_{\mathrm{r}} - \dot{\beta}\right) \\ 0 &= k_{\mathrm{r}}\check{I} - \dot{\beta}\check{R}. \end{aligned}$$

Obviously, the disease free case must always be an equilibrium point, leading to

$$E_0: \quad \left(\check{S}, \check{I}, \check{R}\right) = (1, 0, 0)\,.$$

However, this time, equilibrium may also occur with $\check{I} \neq 0$. With equilibrium infection, the second equation requires

$$\check{S} = \frac{1}{k_{\mathrm{i}}}\left(\dot{\beta} + k_{\mathrm{r}}\right),$$

where we must have

$$0 \leq \check{S} \leq 1 \Rightarrow 0 \leq \frac{1}{k_{\mathrm{i}}}\dot{\beta} + \frac{k_{\mathrm{r}}}{k_{\mathrm{i}}} \leq 1.$$

The third equation requires

$$\check{R} = \frac{k_{\mathrm{r}}}{\dot{\beta}}\check{I}.$$

Finally,

$$\begin{aligned} 1 &= \check{S} + \check{I} + \check{R} \\ &\Downarrow \\ 1 &= \frac{1}{k_{\mathrm{i}}}\dot{\beta} + \frac{k_{\mathrm{r}}}{k_{\mathrm{i}}} + \check{I} + \frac{k_{\mathrm{r}}}{\dot{\beta}}\check{I} \\ &\Downarrow \\ \check{I} &= \frac{1 - \frac{\dot{\beta}}{k_{\mathrm{i}}} - \frac{k_{\mathrm{r}}}{k_{\mathrm{i}}}}{1 + \frac{k_{\mathrm{r}}}{\dot{\beta}}}. \end{aligned}$$

Thus, we find a second, unique equilibrium point

$$E_1: \quad \left(\check{S}, \check{I}, \check{R}\right) = \left(\frac{\dot{\beta} + k_{\mathrm{r}}}{k_{\mathrm{i}}}, \frac{1 - \frac{\dot{\beta}+k_{\mathrm{r}}}{k_{\mathrm{i}}}}{1 + \frac{k_{\mathrm{r}}}{\dot{\beta}}}, \frac{1 - \frac{\dot{\beta}+k_{\mathrm{r}}}{k_{\mathrm{i}}}}{\frac{\dot{\beta}}{k_{\mathrm{r}}} + 1}\right).$$

Observe that in the limit as $\dot{\beta} \to 0$, $E_1 \to \lim_{\dot{\beta}\to 0}\left(\frac{k_{\mathrm{r}}}{k_{\mathrm{i}}}, \frac{1-\frac{k_{\mathrm{r}}}{k_{\mathrm{i}}}}{1+\frac{k_{\mathrm{r}}}{\dot{\beta}}}, 1 - \frac{k_{\mathrm{r}}}{k_{\mathrm{i}}}\right) \to \left(\frac{k_{\mathrm{r}}}{k_{\mathrm{i}}}, 0, 1 - \frac{k_{\mathrm{r}}}{k_{\mathrm{i}}}\right)$, thus $\left(\frac{1}{\mathsf{R}_0}, 0, 1 - \frac{1}{\mathsf{R}_0}\right)$ which is one of possible solutions for the case of no birth/death.

3.2 SIR: condition for infection growth

For infection to be *initiated* from a disease-free situation $E_0 : (1, 0, 0)$, it is common to linearize the model in the disease free equilibrium. Since all numbers are positive, it follows that for all times, $\check{S}(t) \leq \check{S}(0)$. At initial time, we have for the number of infected that

$$\left.\frac{\mathrm{d}\check{I}}{\mathrm{d}t}\right|_{t=0} = \left(k_{\mathrm{i}}\check{S}(0) - k_{\mathrm{r}}\right)\check{I}(0),$$

and we see that the number of infected can only increase if (1) there already are infected individuals (a minimum of 1), and (2) if

$$k_{\mathrm{i}}\check{S}(0) - k_{\mathrm{r}} > 0 \Rightarrow \underbrace{\frac{k_{\mathrm{i}}}{k_{\mathrm{r}}}}_{=\mathsf{R}_0} \cdot \check{S}(0) > 1.$$

For a population that has not been vaccinated, $\check{S}(0) \approx 1$, and for that case, infection growth requires that the basic reproduction number is larger than unity, $\mathsf{R}_0 > 1$. Vice versa, if $\mathsf{R}_0 < 1$, the infection decreases.[27]

More formal methods for studying stability of the disease-free case exist, based on Lyupunov theory (Martcheva 2015). However, here, we will only consider the approach based on linearization in the (near) disease-free case.

3.3 SIR: initial growth rate

If $\check{I}(0)$ is relatively small, $\check{S}(t)$ will be relatively constant for some time $t > 0$, and for some small positive interval in t, we will have $\check{S}(t) \approx \check{S}(0)$. Under that assumption and for no birth/death, the infection equation becomes a linear differential equation,

$$\begin{aligned} \frac{\mathrm{d}\check{I}}{\mathrm{d}t} &\approx \left(k_{\mathrm{i}}\check{S}(0) - k_{\mathrm{r}}\right)\check{I} \\ &\Downarrow \\ \frac{\mathrm{d}\check{I}}{\mathrm{d}t} &\approx \left(\mathsf{R}_0\check{S}(0) - 1\right)k_{\mathrm{r}}\check{I}. \end{aligned}$$

This approximate model has the solution

$$\check{I}(t) = \exp\left(\left(\mathsf{R}_0\check{S}(0) - 1\right)k_{\mathrm{r}}t\right)\check{I}(0),$$

or

$$\ln\check{I} = \left(\mathsf{R}_0\check{S}(0) - 1\right)k_{\mathrm{r}}t + \ln\check{I}(0). \tag{23}$$

Obviously, $I = N \cdot \check{I} \Rightarrow \ln N + \ln\check{I}$ and $\ln I(0) = \ln N + \ln\check{I}(0)$, so $\ln N$ cancels and we have

$$\ln I = \left(\mathsf{R}_0\check{S}(0) - 1\right)k_{\mathrm{r}}t + \ln I(0). \tag{24}$$

Thus, by observing $\ln I$ as a function of t, the initial value is $\ln I(0)$, and the slope is $\left(\mathsf{R}_0\check{S}(0) - 1\right)k_{\mathrm{r}}$.

Example 1. Initial growth rate of measles outbreak

As a case study, consider the boarding school measles outbreak in North England, 1978, with logarithmic scale ordinate axis, Fig. 17.

Observe that for days 3–5, the increase appears linear in the logarithmic scale. We then have

$$\ln I(5) = \left(\mathsf{R}_0\check{S}(3) - 1\right)k_{\mathrm{r}}(5-3) + \ln I(3).$$

Inserting data, where $I(5) = I^{\mathrm{d}}[3] = 227$ while $I(3) = I^{\mathrm{d}}[1] = 25$, while $\check{S}(3) = \frac{763-25}{763} \approx 0.967$, we find

$$\left(\mathsf{R}_0 0.967 - 1\right)k_{\mathrm{r}} = \frac{\ln I(5) - \ln I(3)}{5-3} = \frac{\ln 227 - \ln 25}{2} \approx 1.103.$$

[27] Some authors define R_0 as $\frac{k_{\mathrm{i}}}{k_{\mathrm{r}}}\check{S}(0)$, and talk about a "dynamic" $\mathsf{R}_t = \frac{k_{\mathrm{i}}}{k_{\mathrm{r}}}\check{S}(t)$ which should be less than zero for stability. Normally, this is a too simplistic method to analyze stability of nonlinear systems, but because of the structure of the model, where $\check{S}(t)$ will decrease asymptotically to zero, it may work here. This extended definition of R_0 would explain how R_0 is reduced by vaccination, leaving $\check{S}(0) < 1$, and possibly $\check{S}(0) \ll 1$.

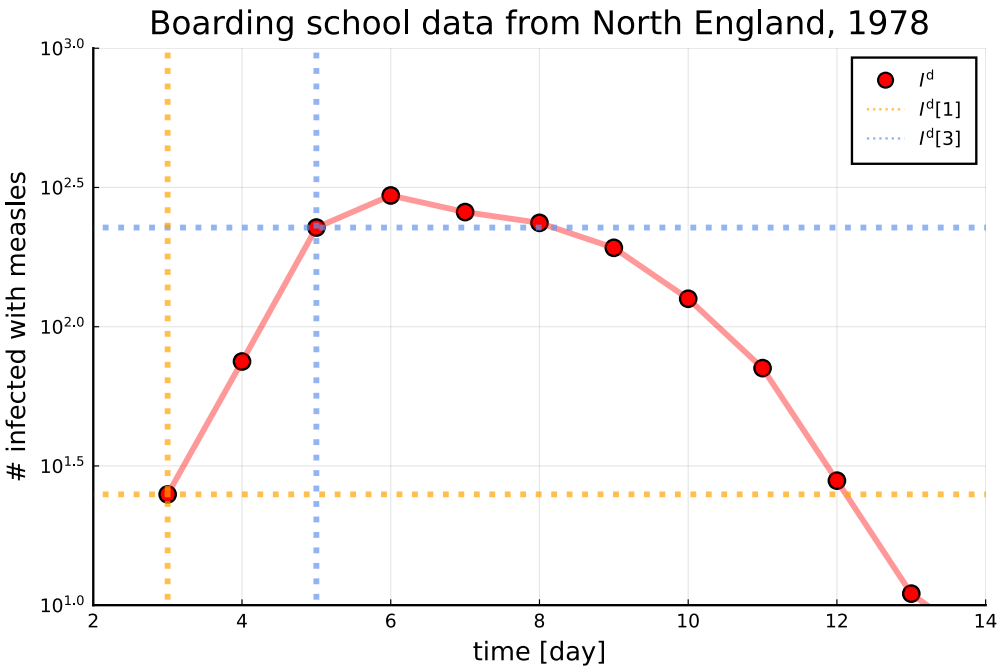

Figure 17: Logarithm of infection development for measles outbreak in North-England, 1978.

Suppose we claim to know that $k_r = 0.3$. Then R_0 is given as

$$\mathsf{R}_0 = \frac{\frac{1.103}{k_r} + 1}{0.967} = \frac{\frac{1.103}{0.3} + 1}{0.967} \approx 4.83.$$

Had we instead used $k_r = 0.15$ which is more in line with the data in Table 1, we would find

$$\mathsf{R}_0 = \frac{\frac{1.103}{0.15} + 1}{0.967} \approx 8.6$$

which would still be lower than expected.

If we have found R_0 this way, we can next find the infection frequency k_i as

$$\mathsf{R}_0 = \frac{k_i}{k_r} \Rightarrow k_i = k_r \mathsf{R}_0 = \begin{cases} 0.3 \cdot 4.83 \approx 1.449, & k_r = 0.3 \\ 0.15 \cdot 8.6 \approx 1.29, & k_r = 0.15. \end{cases}$$

Vice versa, if we instead know k_i and $\check{S}(0)$ in advance, and have found the slope 1.103, we can compute k_r from $k_i \check{S}(0) - k_r = 1.103$ and finally $\mathsf{R}_0 = k_i/k_r$. ▲

The expressions in Eqs. 23 and 24 are probably only good when $S(0) \approx N$, or $\check{S}(0) \approx 1$. When $\check{S}(0) < 1$, $\check{S}(t)$ may change rapidly, and it is unrealistic to assume that $\check{S}(0)$ is constant.

In summary, if we know one of k_r, R_0, or k_i in advance, we can find the missing values of (k_r, R_0, k_i) graphically from the slope of the infection development.

3.4 SIR: doubling time

Another interesting question is: can we say anything about the time T_2 for $\check{I}$ to double its value? This can be found by $T_2 : \check{I}(T_2) = 2\check{I}(0)$. Because $\ln\left(2 \cdot \check{I}(0)\right) = \ln 2 + \ln \check{I}(0)$,

we find from Eq. 23:

$$\begin{aligned}\ln 2 + \ln \check{I}(0) &= \left(\mathsf{R}_0 \check{S}(0) - 1\right) k_\mathrm{r} T_2 + \ln \check{I}(0) \\ &\Downarrow \\ T_2 &= \frac{\ln 2}{\left(\mathsf{R}_0 \check{S}(0) - 1\right) k_\mathrm{r}} \approx \frac{0.69}{\left(\mathsf{R}_0 \check{S}(0) - 1\right) k_\mathrm{r}},\end{aligned}$$

or 0.69 multiplied by the inverse of the slope. This only makes sense if $\mathsf{R}_0 > 1$.

Example 2. Measles outbreak doubling time

From Example 1, we find (with $k_\mathrm{r} = 0.3$) that the doubling time is

$$T_2 \approx \frac{0.69}{(4.83 \cdot 0.967 - 1) \cdot 0.3} \approx 0.63\,\mathrm{d} \approx 15\,\mathrm{h}.$$

Thus, the number of infections double every 15 h. ▲

3.5 SIR: model order reduction$^+$

Consider the ratio of infection change vs. susceptible change, Eq. 9 divided by Eq. 8:

$$\begin{aligned}\frac{\frac{\mathrm{d}\check{I}}{\mathrm{d}t}}{\frac{\mathrm{d}\check{S}}{\mathrm{d}t}} &= \frac{\left(k_\mathrm{i}\check{S} - k_\mathrm{r}\right)\check{I}}{-k_\mathrm{i}\check{I}\check{S}} \\ &\Downarrow \\ \frac{\mathrm{d}\check{I}}{\mathrm{d}\check{S}} &= -1 + \frac{k_\mathrm{r}}{k_\mathrm{i}\check{S}}.\end{aligned}$$

By inserting the expression for R_0, we have

$$\frac{\mathrm{d}\check{I}}{\mathrm{d}\check{S}} = -1 + \frac{1}{\mathsf{R}_0\check{S}}.$$

This expression for $\frac{\mathrm{d}\check{I}}{\mathrm{d}\check{S}}$ can be solved by the method of separation of variables:

$$\begin{aligned}\frac{\mathrm{d}\check{I}}{\mathrm{d}\check{S}} &= -1 + \frac{1}{\mathsf{R}_0\check{S}} \\ &\Downarrow \\ \mathrm{d}\check{I} &= -\mathrm{d}\check{S} + \frac{1}{\mathsf{R}_0}\frac{\mathrm{d}\check{S}}{\check{S}} \\ &\Downarrow \\ \left[\check{I}\right]_{\check{I}(0)}^{\check{I}(t)} &= -\left[\check{S}\right]_{\check{S}(0)}^{\check{S}(t)} + \frac{1}{\mathsf{R}_0}\left[\ln \check{S}\right]_{\check{S}(0)}^{\check{S}(t)}.\end{aligned}$$

It follows that

$$\check{I}(t) = \check{S}(0) + \check{I}(0) - \check{S}(t) + \frac{1}{\mathsf{R}_0}\ln\left(\frac{\check{S}(t)}{\check{S}(0)}\right),$$

where $\check{S}(0)+\check{I}(0)+\check{R}(0)=1$ and $\check{R}(0)$ would represent those that are initially immune. Thus, $\check{S}(0)+\check{I}(0)=1-\check{R}(0)$, or

$$\check{I}(t)=1-\check{R}(0)+\check{S}(t)+\frac{1}{\mathsf{R}_0}\ln\left(\frac{\check{S}(t)}{\check{S}(0)}\right) \tag{25}$$

By inserting Eq. 25 into Eq. 8, we find:

$$\begin{aligned}
&\frac{\mathrm{d}\check{S}}{\mathrm{d}t}=-k_{\mathrm{i}}\check{I}\check{S}\\
&\quad\Downarrow\\
&\frac{\mathrm{d}\check{S}}{\mathrm{d}t}=-k_{\mathrm{i}}\check{S}\left(1-\check{R}(0)+\check{S}(t)+\frac{1}{\mathsf{R}_0}\ln\left(\frac{\check{S}(t)}{\check{S}(0)}\right)\right).
\end{aligned} \tag{26}$$

By solving the first order ODE in Eq. 26, we find $\check{S}(t)$. We can then insert this solution into the algebraic equation for $\check{I}$, Eq. 25, to find $\check{I}(t)$. Finally, we find $\check{R}$ from $\check{R}(t)=1-\check{S}(t)-\check{I}(t)$.

3.6 SIR: maximum infection$^+$

The infection is at its maximum $\check{I}^*$ in time when $\left.\frac{\mathrm{d}\check{I}}{\mathrm{d}t}\right|_{\check{I}^*}=0$ while $\check{I}^*>0$. Using Eq. 9, we find

$$\begin{aligned}
&\frac{\mathrm{d}\check{I}^*}{\mathrm{d}t}=k_{\mathrm{i}}\check{I}^*\check{S}^*-k_{\mathrm{r}}\check{I}^*=0\\
&\quad\Downarrow\\
&0=k_{\mathrm{i}}\check{I}^*\check{S}^*-k_{\mathrm{r}}\check{I}^*\\
&\quad\Downarrow\\
&\check{S}^*=\frac{k_{\mathrm{r}}}{k_{\mathrm{i}}}=\frac{1}{\mathsf{R}_0}.
\end{aligned}$$

Obviously, $\check{S}(t)\leq 1$ which means that $\check{S}^*\leq 1$. This means that the expression is only valid for $\mathsf{R}_0\geq 1$, which makes sense: with $\mathsf{R}_0<1$, there can be no growth in infection and no maximum.

This gives us the value $\check{S}^*$ of $\check{S}$ when $\check{I}$ is at its maximum, but we are more interested $\check{I}^*$. To find $\check{I}^*$, We insert $\check{S}^*$ into Eq. 25:

$$\begin{aligned}
&\check{I}^*=1-\check{R}(0)-\frac{1}{\mathsf{R}_0}+\frac{1}{\mathsf{R}_0}\ln\left(\frac{\check{S}^*}{\check{S}(0)}\right)\\
&\quad\Downarrow\\
&\check{I}^*=1-\check{R}(0)-\frac{1-\ln\left(\frac{1}{\mathsf{R}_0\check{S}(0)}\right)}{\mathsf{R}_0}
\end{aligned}$$

where

$$\ln\left(\frac{1}{\mathsf{R}_0\check{S}(0)}\right)=-\ln\left(\mathsf{R}_0\check{S}(0)\right)=-\ln\mathsf{R}_0-\ln\check{S}(0).$$

Infection can only take place when $\check{I}(0) > 0$, leading to

$$\check{I}^* = 1 - \check{R}(0) - \frac{1 + \ln \mathsf{R}_0 + \ln \check{S}(0)}{\mathsf{R}_0}.$$

In the case of new diseases where $\check{R}(0) = 0$ and $\check{S}(0) \gg \check{I}(0) \Rightarrow \check{S}(0) \approx 1$, we get the simpler expression

$$\check{I}^* = 1 - \frac{1 + \ln \mathsf{R}_0}{\mathsf{R}_0}.$$

To repeat, these expressions for the maximum value of $\check{I}^*$ is valid when $\mathsf{R}_0 \check{S}(0) \geq 1$. Clearly, $\check{I}^*$ increases when R_0 increases, while lowering R_0 through mitigation reduces $\check{I}^*$.

Example 3. Measles outbreak: Use of maximum infection to find R_0

Assuming that model parameters are fixed throughout the measles outbreak, we observe from Fig. 3 or the available data that $\check{I}^* \approx \frac{296}{763} \approx 0.388$, $\check{R}(0) = 0$, while $\check{S}(0) \approx 0.967$. Thus, we have an implicit expression for R_0:

$$\begin{aligned}
\check{I}^* &= 1 - \frac{1 + \ln \mathsf{R}_0 + \ln 0.967}{\mathsf{R}_0} \\
&\Downarrow \\
0.388 &\approx 1 - \frac{1 + \ln \mathsf{R}_0 - 0.034}{\mathsf{R}_0} \\
&\Downarrow \\
\frac{0.966 + \ln \mathsf{R}_0}{\mathsf{R}_0} &\approx 0.612.
\end{aligned}$$

The simplest way to solve this equations for R_0 is probably by issuing command `solve ((0.966+ln(R0))/R0 = 0.612)` to Wolfram Alpha[28], leading to two real solutions:

$$\mathsf{R}_0 \in \{0.525, 3.73\}.$$

Because we have infection growth, we conclude that $\mathsf{R}_0 \approx 3.73$.

By combining this number for R_0 with the growth rate in Example 1, we can re-assess k_r and k_i. We find

$$\begin{aligned}
\mathsf{R}_0 &= \frac{1.103/k_\mathrm{r} + 1}{0.967} \\
&\Downarrow \\
3.73 &= \frac{1.103/k_\mathrm{r} + 1}{0.967} \\
&\Downarrow \\
k_\mathrm{r} &= \frac{1.103}{3.73 \cdot 0.967 - 1} \approx 0.423.
\end{aligned}$$

We then finally find

$$\mathsf{R}_0 = \frac{k_\mathrm{i}}{k_\mathrm{r}} \Rightarrow k_\mathrm{i} = \mathsf{R}_0 k_\mathrm{r} \approx 3.73 \cdot 0.423 \approx 1.58.$$

[28] `https://www.wolframalpha.com/`

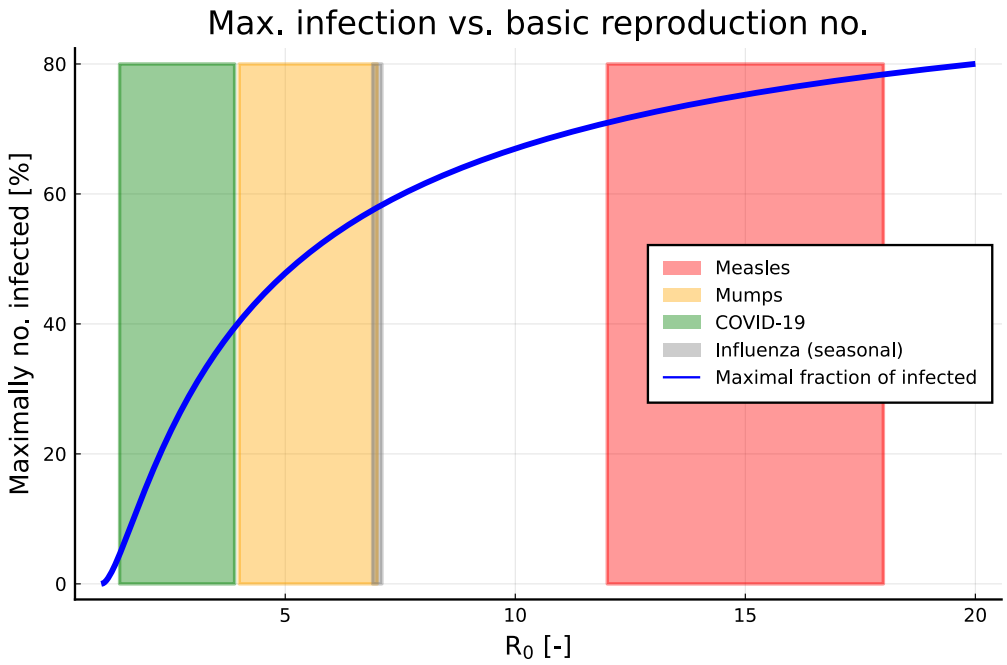

Figure 18: Maximal fraction of infected $\check{I}^*$ [%] vs. basic reproduction number R_0, with range of R_0 for *unmitigated* Measles, for unmitigated Mumps, for COVID-19, and for unmitigated Seasonal Influenza.

These re-assessed numbers $(k_\mathrm{i}, k_\mathrm{r}) = (1.58, 0.423)$ should be compared to those found from model fitting in Section 2.8.3, $(k_\mathrm{i}, k_\mathrm{r}) = (1.81, 0.459)$. ▲

Figure 18 shows the maximal number of infected [%] vs. basic reproduction number, with ranges in R_0 for a few diseases.

Observe that the maximal fraction of infected, $\check{I}^*$, is related to the so-called *herd immunity*: if the fraction of infected reaches $\check{I}^*$ for the unmitigated value of R_0, then the infection spread will die out when lock-down, face mask, social distancing, etc. is canceled. This means that if the unmitigated value of R_0 for measles is ca. 12, then heard immunity is reached when $\check{I}^* \approx 0.7$, or when 70% of the population has been infected, and have recovered = become immune. This fraction of infected to achieve herd immunity obviously is only valid if the population mixes freely so that the compartment model is valid.

3.7 SIR: time until peak infection$^+$

In practice, *mitigation* of infections such as social distancing, lock-down, use of face mask, etc. changes the infection frequency factor k_i, and thereby the basic reproduction number R_0.

Figure 18 shows that mitigation, which changes R_0, will change the maximal infection fraction $\check{I}^*$. This is often described as "flattening the curve", and is important in order to avoid overload on hospitals. However, "flattening" the curve by reducing R_0 leads to a delay in the time until maximum infection is reached. Let us consider the time until the infection peaks.

Table 3: Fractional time $\frac{T_\mathrm{i}^*}{\tau_\mathrm{r}}\left(\mathsf{R}_0;\check{I}\left(0\right)\right)$ until peak infection assuming none are initially immune. T_i^* is time until peak infection, τ_r is the recovery time constant, R_0 is basic reproduction number, and $\check{I}\left(0\right)$ is fractional initial infection.

$\check{I}\left(0\right)\backslash\mathsf{R}_0$	6.6	4.7	2.8	2.4	1.3
1%	1.2	1.6	2.9	3.5	7.8
5%	0.86	1.2	1.9	2.2	2.6
10%	0.72	0.96	1.5	1.6	1.1
30%	0.46	0.56	0.65	0.63	–

The reduced order model in Eq. 26,

$$\frac{\mathrm{d}\check{S}}{\mathrm{d}t}=-k_\mathrm{i}\check{S}\left(1-\check{R}\left(0\right)-\check{S}+\frac{1}{\mathsf{R}_0}\ln\left(\frac{\check{S}}{\check{S}\left(0\right)}\right)\right)$$

is a separable differential equations which can be integrated as

$$\int_{\check{S}(0)}^{\check{S}(t)}\frac{\mathrm{d}\check{S}}{\check{S}\left(1-\check{R}\left(0\right)\right)-\check{S}^2+\frac{\check{S}}{\mathsf{R}_0}\ln\frac{\check{S}}{\check{S}(0)}}=-k_\mathrm{i}\int_0^t\mathrm{d}t=-k_\mathrm{i}t.$$

Unfortunately, there is no known analytic solution to the left hand side integral.

The maximal value $\check{I}^*$ is found when $\check{S}^*=\frac{1}{\mathsf{R}_0}$. Also, k_i depends on R_0 and may vary, so it is better to express k_i as $k_\mathrm{i}=\mathsf{R}_0k_\mathrm{r}=\mathsf{R}_0/\tau_\mathrm{r}$ since k_r (or τ_r) normally would be constant. The time T_i^* taken to reach the maximum infection $\check{I}^*$ can thus be posed as

$$\frac{T_\mathrm{i}^*}{\tau_\mathrm{r}}=\frac{1}{\mathsf{R}_0}\int_{\frac{1}{\mathsf{R}_0}}^{\check{S}(0)}\frac{\mathrm{d}\check{S}}{\check{S}\left(1-\check{R}\left(0\right)\right)-\check{S}^2+\frac{\check{S}}{\mathsf{R}_0}\ln\frac{\check{S}}{\check{S}(0)}}$$

where

$$\check{S}\left(0\right)+\check{I}\left(0\right)+\check{R}\left(0\right)=1.$$

It follows that the fraction $\frac{T_\mathrm{i}^*}{\tau_\mathrm{r}}$ depends on both R_0, $\check{S}\left(0\right)$, and $\check{R}\left(0\right)$. This fraction can be found by numeric *quadrature*, and we should expect that a solution can only be found when $\mathsf{R}_0\check{S}\left(0\right)>1$.

Figure 19 shows the fractional time $\frac{T_\mathrm{i}^*}{\tau_\mathrm{r}}$ until peak infection vs. basic reproduction number when $\check{R}\left(0\right)\equiv 0$. Thus, zero immune individuals has been assumed, i.e., $\check{S}\left(0\right)=1-\check{I}\left(0\right)$, and $\check{I}\left(0\right)$ has been used as a parameter.

Table 3 shows numerical values for some relevant basic reproduction numbers R_0 for COVID-19, for a number of initial infection fractions. In Table 3, observe that for $\check{I}\left(0\right)=30\%$ (0.3) and $\mathsf{R}_0=1.3$, $\mathsf{R}_0\check{S}\left(0\right)=1.3\cdot(1-0.3)=0.91<1$, and the infection will die out without any peak. It is straightforward to find similar numbers with an initial immunity fraction.

Example 4. SIR measles case

The case of constant number of individuals N is considered, with initially recovered fraction $R\left(3\right)=0$ (i.e., no-one is immune) and recovery frequency factor $k_\mathrm{r}\approx 0.459$ or recovery time constant $\tau_\mathrm{r}=1/k_\mathrm{r}\approx 2.18\,\mathrm{d}$. Basic reproduction number is either $\mathsf{R}_0=3.9$

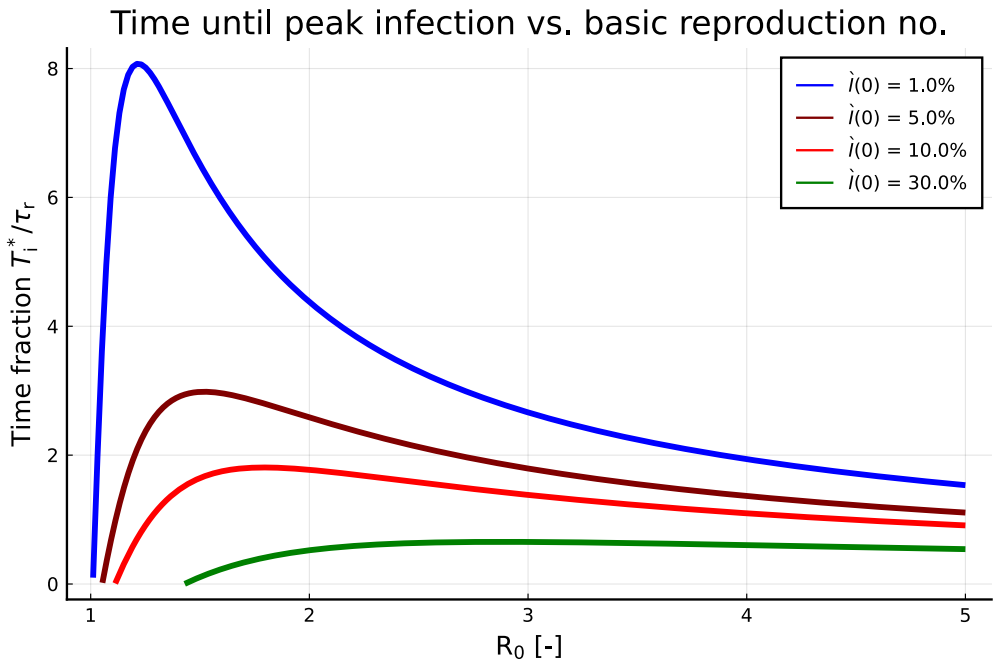

Figure 19: Fractional time $\frac{T_i^*}{\tau_r}$ until peak infection vs. basic reproduction number R_0, with $\check{I}(0)$ as parameter.

(fitted to data) or $R_0 = 2.6$ (hypothetical case). Initial number of infected is either $I(3) = 25$ (1%).

Figure 20 illustrates the difference in behavior with $R_0 = 3.9$ vs. $R_0 = 2.6$, and includes computed I^* value.

▲

From Example 4, observe that "flattening the curve", i.e., decreasing the infection frequency factor k_i and thus reducing R_0, implies delaying the time where the infection peaks.

3.8 Generalizing the Basic Reproduction Number$^+$

We have so far considered the basic reproduction number R_0 for the SIR model. In Section 2.4, we just claimed the expression $R_0 = \frac{k_i}{k_r}$, while in Section 3.2, we saw that $R_0 \cdot \check{S}(0) > 1$ is a condition for infection growth, and with $\check{S}(0) \equiv 1$, initial infection will grow if $R_0 > 0$.

We need to generalize the idea of basic reproduction number for more complex models. The basic idea is to consider the stability of the *disease-free, initial state*. Disease-free implies that $\check{I}(0) \equiv 0$. If there is an immune fraction $\check{R}(0)$, then the initial fraction of susceptibles is $\check{S}(0) = 1 - \check{R}(0)$. All other compartments have zero fraction. If we neglect immigration and emmigration, and set N constant, we have

$$\frac{\mathrm{d}\check{X}}{\mathrm{d}t} = \nu^{\mathsf{T}} r$$

where $r = r\left(\check{X}\right)$. If $\check{X} = \left(\check{S}, \check{I}, \ldots, \check{R}\right)$ and $\check{R}(0) = 0$, then $\check{X}(0) = (1, 0, \ldots, 0)$. Stability

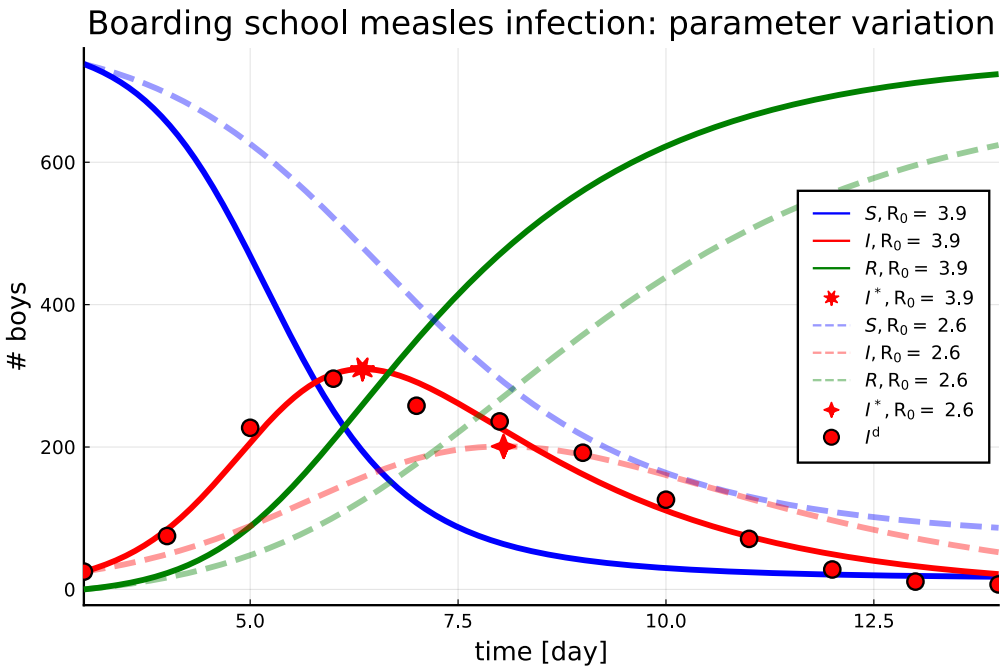

Figure 20: Measles: comparison of $\mathsf{R}_0 = 3.9$ (fitted data) vs. $\mathsf{R}_0 = 2.6$ (hypothetical). In both cases, $I(3) = 25$, $R(3) = 0$, $k_\mathrm{r} = 0.459\,\mathrm{d}^{-1}$ (or $\tau_\mathrm{r} = 2.18\,\mathrm{d}$).

can then be assessed by the eigenvalues of the Jacobian $J_{\check{X}(0)}$ given by

$$J_{\check{X}(0)} = \frac{\partial}{\partial \check{X}} \left(\nu^\mathsf{T} r\left(\check{X}\right)\right)\Big|_{\check{X}(0)}.$$

If all eigenvalues of $J_{\check{X}(0)}$ have negative real part, then the system is stable, and there can not be infection growth. If, however, at least one eigenvalue have positive real part, then infection can grow.

This means that for a well tuned model, we can find the initial Jacobian of the model and check the eigenvalues to find whether infection can grow or not. For the SIR model, the single eigenvalue is

$$\begin{aligned} \lambda &= -\left(k_\mathrm{i}\check{S}(0) - k_\mathrm{r}\right) = -k_\mathrm{r}\left(\mathsf{R}_0\check{S}(0) - 1\right) \\ &\Downarrow \\ 0 &= \lambda + k_\mathrm{r}\left(\mathsf{R}_0\check{S}(0) - 1\right), \end{aligned} \tag{27}$$

where Eq. 27 is the characteristic equation for this simple case. Although the sign of the real part of eigenvalues provides the *foundation* for finding the basic reproduction number, eigenvalues do not directly provide the basic reproduction number.

In general, the basic reproduction number should (Martcheva 2015):

- Be non-negative for non negative parameter values,
- Be zero if there is no transmission,
- Be interpretable as the number of secondary infections.

Finding symbolic expressions for the eigenvalues is non-trivial. Alternatively, one can base the stability requirement on the characteristic equation of the Jacobian,

$$\left|J_{\check{X}(0)} - \lambda I\right| = 0 \Rightarrow \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n = 0.$$

It is a necessary (but not sufficient) requirement that $\forall i : a_i > 0$ to ensure that all eigenvalues have negative real part. To find sufficient requirements, the Routh-Hurwitz criterion often provides the answer.

Theorem 1. *Routh-Hurwitz theorem*

We consider a matrix $A \in \mathbb{R}^{n\times n}$ with characteristic polynomial $\varphi(\lambda)$, and characteristic equation

$$\varphi(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n = 0,$$

and introduce n Hurwitz matrices H_j, $j \in \{1, \ldots, n\}$ defined as

$$H_j = \begin{pmatrix} a_1 & 1 & 0 & 0 & \cdots & 0 \\ a_3 & a_2 & a_1 & 1 & \cdots & 0 \\ a_5 & a_4 & a_3 & a_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \\ 0 & 0 & 0 & 0 & 0 & a_j \end{pmatrix}$$

where $a_j = 0$ if $j < 0$, $a_0 = 1$, and $a_j = 0$ if $j > n$.

Then all eigenvalues of A have negative real parts if

$$\det H_j > 0, \quad j \in \{1, \ldots, n\}.$$

▲

As an example, let us consider the SEIR model.

Example 5. Basic Reproduction Number for SIR model with birth-death

The model is given in Eqs. 20–20, leading to Jacobian

$$J_{\check{X}} = \begin{pmatrix} -1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} k_{\mathrm{i}}\check{I} & k_{\mathrm{i}}\check{S} & 0 \\ 0 & k_{\mathrm{r}} & 0 \end{pmatrix} - \dot{\beta}\mathsf{I}$$

where $\check{X} = \left(\check{S}, \check{I}, \check{R}\right)$ and where I is the identity matrix. Multiplying the expression out leads to

$$J_{\check{X}} = \begin{pmatrix} -k_{\mathrm{i}}\check{I} - \dot{\beta} & -k_{\mathrm{i}}\check{S} & 0 \\ k_{\mathrm{i}}\check{I} & k_{\mathrm{i}}\check{S} - k_{\mathrm{r}} - \dot{\beta} & 0 \\ 0 & k_{\mathrm{r}} & -\dot{\beta} \end{pmatrix},$$

or with $\check{X}(0) = (1, 0, 0)$ as the disease-free case:

$$J_{\check{X}(0)} = \begin{pmatrix} -\dot{\beta} & -k_{\mathrm{i}} & 0 \\ 0 & k_{\mathrm{i}} - k_{\mathrm{r}} - \dot{\beta} & 0 \\ 0 & k_{\mathrm{r}} & -\dot{\beta} \end{pmatrix}.$$

Here, the characteristic polynomial is given as

$$\varphi(\lambda) = -\left(\dot{\beta} + \lambda\right)^2 \left(\lambda + \dot{\beta} + k_{\mathrm{r}} - k_{\mathrm{i}}\right).$$

It follows that two eigenvalues are located at $\lambda = -\dot{\beta}$ and are negative ("stable"), while the third eigenvalue is $\lambda = -\dot{\beta} - k_{\mathrm{r}} + k_{\mathrm{i}}$ and is negative when

$$\begin{gathered} -\dot{\beta} - k_{\mathrm{r}} + k_{\mathrm{i}} < 0 \\ \Downarrow \\ k_{\mathrm{i}} - \dot{\beta} < k_{\mathrm{r}} \\ \Downarrow \\ \frac{k_{\mathrm{i}} - \dot{\beta}}{k_{\mathrm{r}}} < 1. \end{gathered}$$

It thus makes sense to define $\mathsf{R}_0 = \frac{k_{\mathrm{i}} - \dot{\beta}}{k_{\mathrm{r}}}$ for this case. ▲

Example 6. Basic Reproduction Number for SEIR model

The model is

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} \check{S} \\ \check{E} \\ \check{I} \\ \check{R} \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} k_{\mathrm{i}} \check{I} \check{S} \\ k_{\mathrm{e}} \check{E} \\ k_{\mathrm{r}} \check{I} \end{pmatrix}$$

leading to Jacobian

$$J_{\check{X}} = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} k_{\mathrm{i}} \check{I} & 0 & k_{\mathrm{i}} \check{S} & 0 \\ 0 & k_{\mathrm{e}} & 0 & 0 \\ 0 & 0 & k_{\mathrm{r}} & 0 \end{pmatrix}.$$

At the disease-free, initial operating point with no-one initially immune, $\check{X}(0) = (1, 0, 0, 0)$, thus

$$J_{\check{X}(0)} = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & k_{\mathrm{i}} & 0 \\ 0 & k_{\mathrm{e}} & 0 & 0 \\ 0 & 0 & k_{\mathrm{r}} & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & -k_{\mathrm{i}} & 0 \\ 0 & -k_{\mathrm{e}} & k_{\mathrm{i}} & 0 \\ 0 & k_{\mathrm{e}} & -k_{\mathrm{r}} & 0 \\ 0 & 0 & k_{\mathrm{r}} & 0 \end{pmatrix}.$$

Clearly, because $J_{\check{X}(0)}$ has two zero columns, its maximal rank is 2, i.e., it must have two eigenvalues at zero. The characteristic polynomial is

$$\varphi(\lambda) = \lambda^2 \left(\lambda^2 + \lambda (k_{\mathrm{e}} + k_{\mathrm{r}}) + k_{\mathrm{e}} (k_{\mathrm{r}} - k_{\mathrm{i}}) \right).$$

Here, we only consider the sub-polynomial that can give instability. Thus, we consider

$$\varphi_1(s) = \lambda^2 + \lambda (k_{\mathrm{e}} + k_{\mathrm{r}}) + k_{\mathrm{e}} (k_{\mathrm{r}} - k_{\mathrm{i}})$$

with Hurwitz matrices

$$H_1 = (k_{\mathrm{e}} + k_{\mathrm{r}})$$

and

$$H_2 = \begin{pmatrix} k_{\mathrm{e}} + k_{\mathrm{r}} & 1 \\ 0 & k_{\mathrm{e}} (k_{\mathrm{r}} - k_{\mathrm{i}}) \end{pmatrix}.$$

Here, clearly $\det H_1 > 0$. Next, we find

$$\begin{aligned}
\det H_2 &= k_{\mathrm{e}}\,(k_{\mathrm{r}} - k_{\mathrm{i}})\,(k_{\mathrm{e}} + k_{\mathrm{r}}) > 0 \\
&\Downarrow \\
0 &< k_{\mathrm{r}} - k_{\mathrm{i}} \\
&\Downarrow \\
0 &< 1 - \frac{k_{\mathrm{i}}}{k_{\mathrm{r}}} \\
&\Downarrow \\
\frac{k_{\mathrm{i}}}{k_{\mathrm{r}}} &< 1.
\end{aligned}$$

Thus, introducing $\mathsf{R}_0 = \frac{k_{\mathrm{i}}}{k_{\mathrm{r}}}$, infection is stopped if $\mathsf{R}_0 < 1$, while infection grows if $\mathsf{R}_0 > 1$. In other words: we find the same expression for the basic reproduction number for the SEIR model as for the SIR model. ▲

In general, the basic reproduction number will be different for different models. As an example, including birth and death rates in the models will lead to different basic reproduction number for the SIR model and the SEIR model, see Keeling & Rohani (2008). See, e.g., Keeling & Rohani (2008) and Martcheva (2015) for more details on how to find an expression for R_0 for various models.

4 Case study: models for COVID-19 spread

4.1 Covid-19 data

Appendix A.1 suggests possible open sources for Covid-19 data. These data provide cumulative numbers of confirmed infected, cumulative numbers of dead, and cumulative numbers of recovered. Because some countries do not register recovered, the cumulative numbers of recovered are of varying quality.

Figure 21 shows the percentage of dead (ordinate) vs. infected (abscissa) for a number of countries.

A notable observation from Fig. 21 is that northern countries (Europe + USA) have a marked change of slope; the change of slope appears to take place ca. 1-2 week of June.[29] The other countries in the figure appear to have a more constant slope (South America, Asia). Some possible reasons why there is a change of slope could be:

- Improved health care which happened to "kick-in" in early June,
- Change of how deaths are classified,
- Mutations of the virus (unlikely?),
- Mitigation effects — hand hygiene, social distancing, quarantining, face masks, closed bars, etc.
- Seasonal effects (northern countries):

[29] Time is the parameter along the curves.

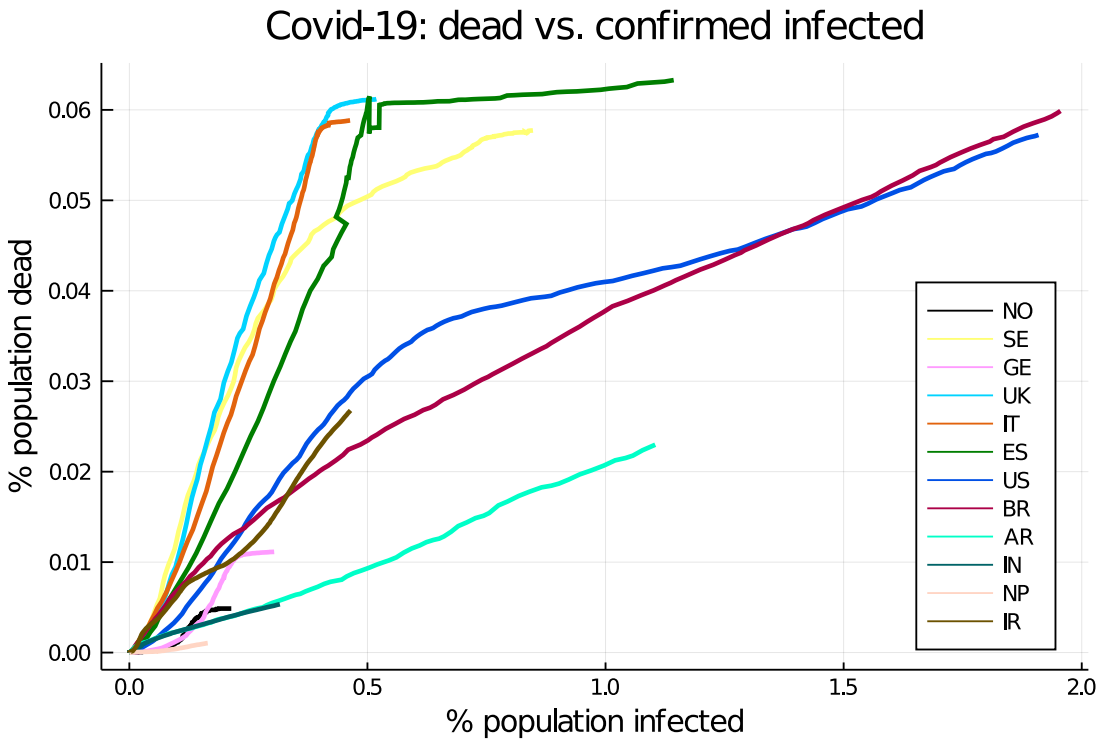

Figure 21: Numbers of deaths per capita (in %) vs. numbers of confirmed infected for a number of countries. Countries are indicated by their internet country code suffix, i.e., NO = Norway, SE = Sweden, etc. The countries are grouped with European countries (NO–ES), American countries (US–AR), and Asian countries (IN–IR). Many other countries could have been included in the comparison, but often have few registered infectees.

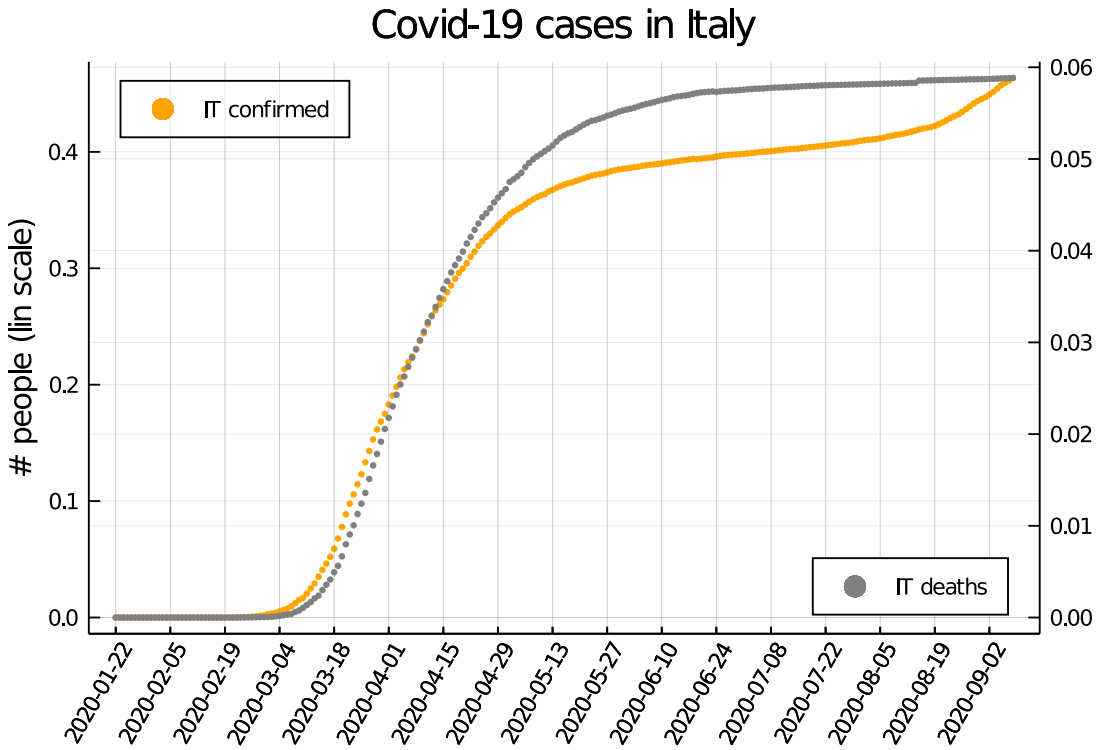

Figure 22: Cumulative number of confirmed infected in Italy vs. time (orange; left ordinate axis), and cumulative number of dead (grey; right ordinate axis).

  - Increased humidity causes aerosols/saliva droplets to travel shorter,
  - Stronger solar irradiation kills virus faster on surfaces,
  - Build-up of D vitamin in plasma due to UV rays,
  - Staying outdoors

- Other causes

Here, we will focus on infection and infection models. Dynamic models have been published for countries South Korea, Italy, and Spain; here we will focus on Italy and Spain. Thus, Fig. 22 shows cumulative number of infected and dead in Italy.

Figure 23 shows cumulative number of infected and dead in Spain.

4.2 Extended SEIR mechanism

One notable observation about covid-19 infection is the distinction between those that are seriously *infected*, and the so-called *asymptomatic* infected. Those that suspect they are infected typically do a test[30], and are classified as non-infected or confirmed infected. We will use symbol C to refer to *confirmed* infected; these will be assumed to be put in isolation and are thus hindered for further spreading.

Those that are only lightly infected may still infect others, but are so little affected by the infection that they do not test themselves. These are the asymptomatic infected,

[30]This is only partially true: in some countries with poor governmental testing, it has been reported that private actors charge in the range of $10-$100 for a test, which may be too expensive in low-income countries.

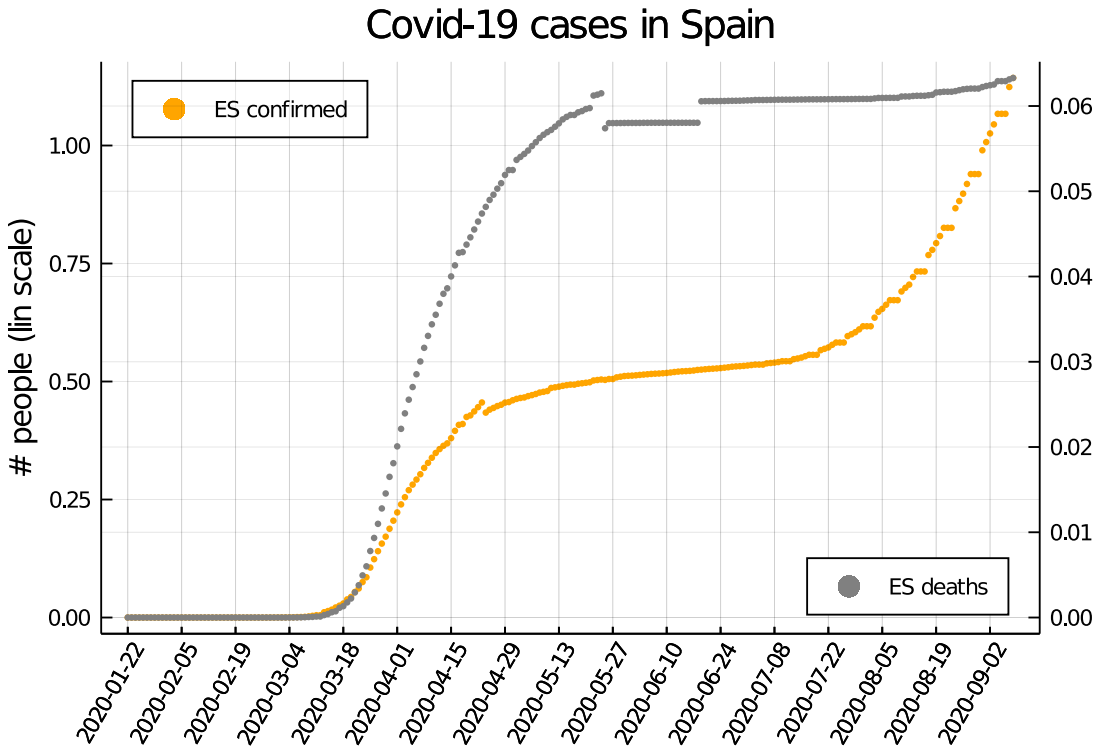

Figure 23: Cumulative number of confirmed infected in Spain vs. time (orange; left ordinate axis), and cumulative number of dead (grey; right ordinate axis).

it is these that continue to infect others; we will use symbol A to refer to *asymptomatic* infected.

An extended SEIR model was proposed in Liu, Magal, Seydi & Webb (2020) with the following mechanism[31]:

$$\begin{aligned}
\mathcal{E}_{\mathrm{i}} &: \mathrm{S} \overset{\mathrm{I+A}}{\to}{}^{k_{\mathrm{i}}} \mathrm{E}, & r_{\mathrm{i}} &= k_{\mathrm{i}} \left(\check{I} + \check{A}\right) \check{S} \\
\mathcal{E}_{\mathrm{e}} &: \mathrm{E} \to^{k_{\mathrm{e}}} \mathrm{I}, & r_{\mathrm{e}} &= k_{\mathrm{e}} \check{E} \\
\mathcal{E}_{\mathrm{c}} &: \mathrm{I} \to^{k_{\mathrm{c}}} \mathrm{C}, & r_{\mathrm{c}} &= k_{\mathrm{c}} \check{I} \\
\mathcal{E}_{\mathrm{a}} &: \mathrm{I} \to^{k_{\mathrm{a}}} \mathrm{A}, & r_{\mathrm{a}} &= k_{\mathrm{a}} \check{I} \\
\mathcal{E}_{\mathrm{cr}} &: \mathrm{C} \to^{k_{\mathrm{r}}} \mathrm{R}, & r_{\mathrm{cr}} &= k_{\mathrm{r}} \check{C} \\
\mathcal{E}_{\mathrm{ar}} &: \mathrm{A} \to^{k_{\mathrm{r}}} \mathrm{R}, & r_{\mathrm{ar}} &= k_{\mathrm{r}} \check{A},
\end{aligned}$$

where we can introduce

$$\begin{aligned}
k_{\mathrm{ca}} &= k_{\mathrm{c}} + k_{\mathrm{a}} \\
k_{\mathrm{c}} &= \eta k_{\mathrm{ca}}.
\end{aligned}$$

Thus, specifying k_{ca} and η, we can find

$$\begin{aligned}
k_{\mathrm{c}} &= \eta k_{\mathrm{ca}} \\
k_{\mathrm{a}} &= (1 - \eta)\, k_{\mathrm{ca}}.
\end{aligned}$$

[31] Here, the notation has been changed to be more in line with the notation of course FM1015 Modelling of Dynamic systems.

The above, proposed mechanism implies that *susceptibles* S are exposed by some "pre-infected" I and the non-quarantined *asymptomatics* A leading to the pre-*infected* phase I. These pre-infected then either become more seriously infected and are *confirmed* infected C, or stay as *asymptomatic* A. Finally, the confirmed infected and the asymptomatic end up in the *recovered* population R (which includes those who die).

The model assumes that the total population is fixed, and has been fitted to data for Italy and Spain, and was published in mid April 2020.[32] In the model fitting, the cumulative number of confirmed infected are used as observation y, i.e.,

$$y(t) = \int_{t_0}^{t} r_{\mathrm{c}}(\tau)\,\mathrm{d}\tau + y(t_0) = k_{\mathrm{c}} \int_{t_0}^{t} \check{I}(\tau)\,\mathrm{d}\tau + y(t_0)$$

where $t \geq t_0$. The model takes into account the effect of mitigation through modeling a time varying infection time k_{i}. Specifically, the model proposes that

$$k_{\mathrm{i}}(t) = \begin{cases} k_{\mathrm{i}}^0, & 0 \leq t \leq T_1 \\ k_{\mathrm{i}}^0 \exp\left(-\frac{t-t_1}{T_{\mathrm{i}}}\right), & t > T_1 \end{cases} \tag{28}$$

where T_1 is the time instance when some mitigation policy takes effect, and T_{i} is the time constant of how the mitigation policy changes k_{i}.

4.3 Mitigation policy response

Equation 28 is an attempt to describe the effect of mitigation. However, the expression is questionable:

- As time goes to infinity ($t \to \infty$), $k_{\mathrm{i}} \to 0$, which essentially presumes that the mitigation is perfect for all future.

A more realistic model would consider some mitigation policy $u(t)$, and propose something like

$$\frac{\mathrm{d}x}{\mathrm{d}t} = f(x, u) \tag{29}$$

$$k_{\mathrm{i}} = g(x, u) \tag{30}$$

where u is the mitigation policy and x is some state which describes the dynamics of the change of k_{i}, and may include people's inertia towards taking measures into use, and the tendency that people get tired of the measures and want to get back to normal life.

Figure 24 indicates two mitigating measures in Norway, mask use and social distancing, as given by `https://covid19.healthdata.org/global?view=total-deaths&tab=trend`. It is important to make a distinction between the mitigation *policy* $u(t)$, and the response in behavior $x(t)$ in the population. Figure 24 displays the population *response*. It is not clear whether the responses in Fig. 24 stem from *observations* or from a model fit.

For Norway, the policy was a strict lock-down ca. March 12, 2020, where people were told to work from home, where restaurants and bars closed, and schools and universities closed. There was a gradual opening of primary schools ca. end of April–beginning of

[32] The model has also been fitted to data from South Korea, but we will neglect tht case here. A correction of one expression was received from the last author in mid September 2020.

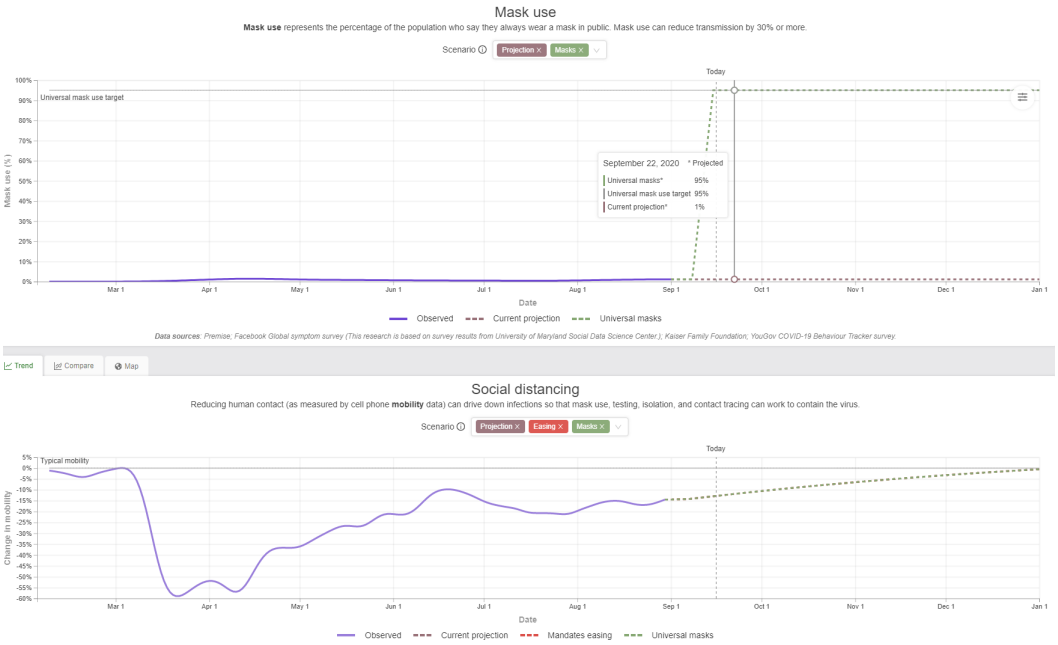

Figure 24: Mask use and social distancing in Norway, according to model of `https://covid19.healthdata.org/global?view=total-deaths&tab=trend`.

May 2020. Universities opened partially in late May 2020. Most things were slow until Midsummer 2020, when vacation started. With strong restrictions on going out of Norway for vacation, most people stayed in Norway. Some infection has been traced back to the few who went on vacation abroad in mid July 2020, and also some infection has been traced to the few tourists that came to Norway in July. Some infection has been traced to a gradual increase in partying, which has taken place partially against the advice of the authorities.

It is difficult to quantity the mitigation *policy* u. If $u = 1$ implies no policy/back to pre covid-19, and $u = 0$ implies no social contact/infection possibility, then perhaps one could say that for Norway,

$$u(t) = \begin{cases} 1, & \text{prior to March 12} \\ 0.45, & [\text{March 12}, \text{April 20}] \\ 0.65, & [\text{April 21}, \text{May 20}] \\ 0.85, & [\text{May 21}, \text{June 23}] \\ 0.8, & [\text{June 24}, \text{August 1}] \\ 0.85, & [\text{August 1}, \text{September 15}] . \end{cases} \tag{31}$$

This is a very qualitative description, and the numbers are taken from some very coarse "guesstimate" of apparent steady levels in the Social distancing panel of Fig. 24 as $u = \frac{100+\text{change in mobility}}{100}$. Face mask is not commonly used in Norway so far, except for being recommended in the main cities and when using public transport.

Similar graphs for Italy are given in Fig. 25. Similar graphs for Spain are given in Fig. 26.

To properly form a model such as the generic model in Eqs. 29 and 30, it would be necessary to have some sort of registration of governmental mitigation policy $u(t)$ similar to Eq. 31, as well as *observations* of mitigation response.

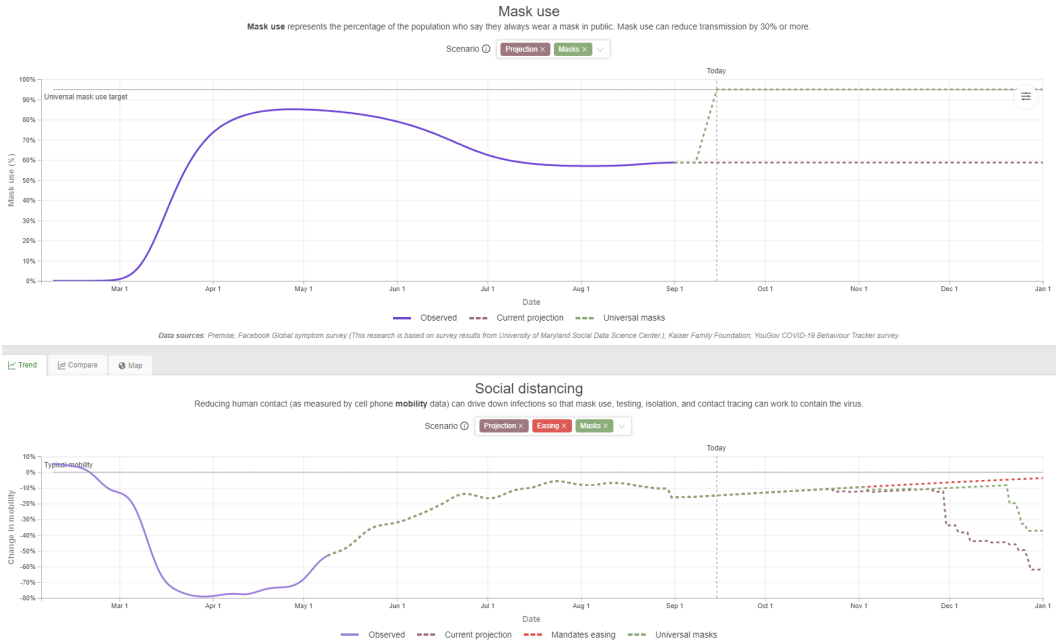

Figure 25: Mask use and social distancing in Italy, according to model of `https://covid19.healthdata.org/global?view=total-deaths&tab=trend`.

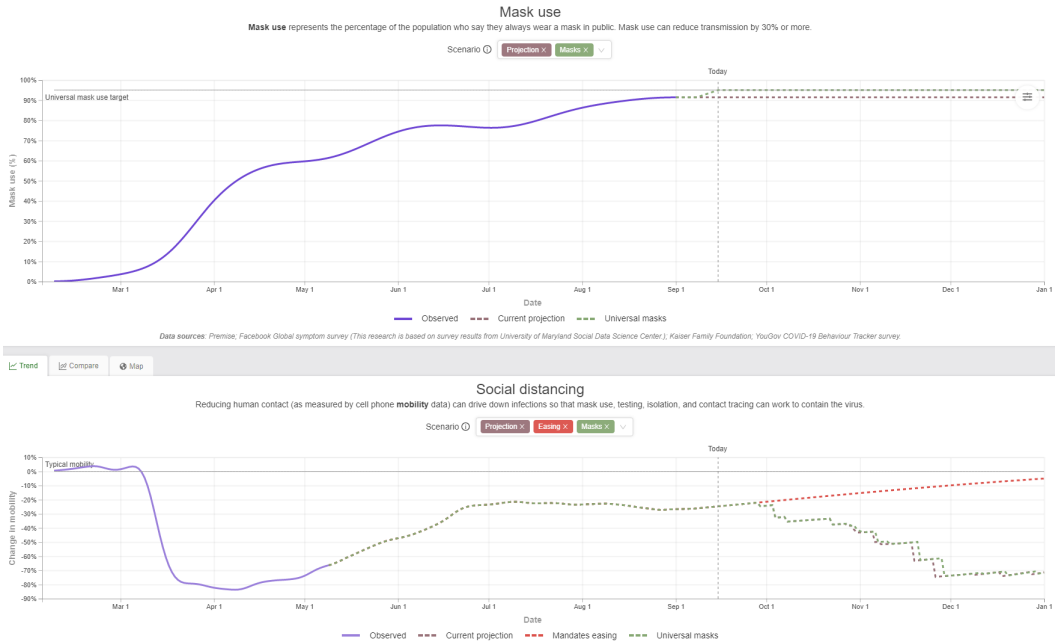

Figure 26: Mask use and social distancing in Spain, according to model of `https://covid19.healthdata.org/global?view=total-deaths&tab=trend`.

Assuming that Eq. 31 gives a correct description for Norway, and that the social distancing plot in Fig. 24 represents an observation of the true social distancing[33] (e.g., through tracking mobile phone data), it would be possible to fit some dynamic model from u_{sd} to state x_{sd}, assuming, say a first order, linear model. Then one could relate k_{i} to x_{sd}.

From Liu, Magal & Webb (2020), it is suggested that $\frac{1}{T_{\mathrm{i}}}$ varies from 0.08 to 0.6 for a number of countries, in other words that the time constant T_{i} for inertia to mitigation policy varies from 1.7 to 12.5 days. Considering the data from Norway in addition, a time constant of $T_{\mathrm{i}} \approx 10\,\mathrm{d}$ could be suggested as an initial guess.

The following simplified version of Eqs. 29 and 30 could be postulated:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \frac{1}{T_{\mathrm{i}}}\left(u - x\right) \tag{32}$$

$$k_{\mathrm{i}} = k_{\mathrm{i}}^{0} x. \tag{33}$$

With this model, in the normal, pre covid-19 situation with $u = 1$, x will asymptotically approach $x = 1$, and $k_{\mathrm{i}} \to k_{\mathrm{i}}^{0}$. As u becomes smaller and smaller due to a low social contact mitigation policy, x will asymptotically approach a smaller and smaller value until it reaches value $x = 0$ for zero social contact. For that case, $k_{\mathrm{i}} \to 0$, which means that the reaction rate $\propto k_{\mathrm{i}}$ approaches zero.

Even more realistic models would take into account that mitigation measures may have a varying effect, in that people may tend to ignore advice such as social distancing over time.

5 Project tasks

5.1 Problems

Consider the following problems.

Problem 1. Getting familiar with infection models

In order to get familiar with infection models, start by studying the SIR model.

1. Implement the SIR model in your modeling language of choice.
2. Use model parameters for the measles outbreak in Northern England, Section 2.8, and simulate the SIR model.
3. Compare the results of your simulation with the experimental data provided in Fig. 5/Table 2. ▲

Problem 2. Covid-19 model for Italy

In this problem, implement a dynamic model based on the extended SEIR mechanism of Section 4.2.

1. Formulate a dynamic model for Italy, with no immigration/emigration, and with model parameters as given in Appendix A.2 and operating conditions as in Appendix A.3. [**NOTE**: the model parameters in Appendix A.2 and the mitigation response as in Eq. 28 are based on data from Italy until the first half of April 2020.]

[33] It is not 100% clear what is meant with a "change in mobility" by, say, 60%.

2. Implement the model and simulate it. Verify the model implementation with comparison to the time evolution of cumulative infected as in Fig. 22. How does the model compare with confirmed data?

3. Would it be possible to simplify the model by eliminating categories/compartments for E and/or I? [**Hint**: it may be necessary with a slight tuning of the parameters given in Appendix A.2.]

▲

Problem 3. Expanding Covid-19 model for Italy with mitigation model

Expand the model in Problem 2 with a more realistic mitigation model as in Eqs. 32–33.

1. Suggest a time evolution or u for Italy, similar to the one proposed for Norway in Eq. 31.

2. Formulate the modified model with mitigation description, implement the model, and verify the implementation.

3. Vary the formulation of u and/or T_{i} to see if you get decent similarity between model and data for Italy as in Fig. 22. ▲

Problem 4. Combined model for Italy and Spain, with population interaction

Based on your experience from Problems 2–3:

1. Modify the model for Italy to allow for immigration and emigration, where the total population remains constant. Assume "mixed compartment" for the emigration term.

2. Develop a similar model for Spain which allows for immigration and emigration.

3. Combine the model for Italy and Spain so that immigration into Italy equals emigration out of Spain, and emigration out of Italy equals immigration into Spain.

4. With the combined model for Italy and Spain, do some experiments with:

   (a) Variation in immigration/emigration rate, and

   (b) Variation in the mitigation *policy*.

   Comment on the findings of the experiments. ▲

Problem 5. Quenching the covid-19 epidemics

Two possible problems of overcoming the covid-19 problem can be considered: achieving herd immunity, and immunity by vaccination.

Herd immunity is achieved by removing any mitigating policy, and let the infection run its course until so many have been infected that there are too few susceptibles remaining to uphold the epidemi.

Vaccination works by artificially transferring susceptibles to recovered by vaccination.

1. *Herd immunity.* Set mitigating policy to $u \equiv 1$. This also implies that k_i must be increased to the frequency factor of the unmitigated case, typically with $\mathsf{R}_0 \approx 4$ or so.

   (a) Simulate the system until herd immunity (for, say, Italy), and find what fraction of the population has been infected.

   (b) Find how long time it takes to reach herd immunity.

   (c) What is a serious negative side effect of trying to reach herd immunity?

2. *Vaccination.* Assume that every remaining susceptible person is vaccinated, and assume a 50% efficiency in the vaccination.

   (a) Simulate the system with and without mitigating policy ($0 \leq u < 0$, and $u \equiv 1$), and observe the spread of infection. ▲

5.2 Documentation

Write a partial report and submit it in the Canvas group folder within the deadline. Write a technical report on the project and results (single file, maximum 15 pages + 1 cover page, PDF format) and submit the report in the Canvas group folder within the deadline. Prepare a "PowerPoint" presentation (=use PowerPoint or other tools) of 7 min for the oral presentation.

A Data, nominal parameters, and operating conditions

A.1 Data for covid-19

It is difficult to find extensive sets of data for covid-19 infections. Web page `https://www.worldometers.info/coronavirus` provides perhaps the most up-to-date tabular and graphical data, but access to numeric timeseries data is expensive. The page does provide up-to-date populations for each country, though. It is also possible to rip data from the graphical presentations; tool Webplotdigitizer `https://automeris.io/WebPlotDigitizer` is recommended — this approach does lead to some inaccuracies. Another useful web page is `https://covid19.healthdata.org/global?view=total-deaths&tab=trend` — this page mainly provides model predictions using advanced models, though. However, it may be possible to extract some information about mitigation policies such as social distancing, use of face masks, etc. Again, the data are only available as graphical presentations.

One of the few web pages that provide free, updated data sets in the form of `*.csv` files[34], is run by Johns Hopkins University. To find the data, google `johns hopkins coronavirus data github`, and choose url `https://github.com/octonion/COVID-19`. Choose folder `csse_covid_19_data`, and then choose `csse_covid_19_time_series`. Next, click on `time_series_covid19_confirmed_global.csv`. This leads to a tabular view of confirmed cases. The following Julia code downloads the `*.csv` files:

[34] csv = Comma Separated Values, a standard file format that can be opened, e.g., in Excel, or imported into mainstream computer languages such as MATLAB, Python, and Julia.

Table 4: Nominal overall parameters for reformulation of extended SEIR model in Liu, Magal & Webb (2020).

Parameter	Value Italy	Value Spain	Relations to original model
N	60 500 000	46 700 000	Populations in original paper
k_{i}^0	$0.25168\,\mathrm{d}^{-1}$	$0.33204\,\mathrm{d}^{-1}$	$k_{\mathrm{i}}^0 = N \cdot \tau_0$; τ_0 in original paper
k_{e}	$1\,\mathrm{d}^{-1}$	$1\,\mathrm{d}^{-1}$	α in original paper
k_{ca}	$1/7\,\mathrm{d}^{-1}$	$1/7\,\mathrm{d}^{-1}$	ν in original paper
η	0.4	0.4	f in original paper
k_{c}	$k_{\mathrm{ca}}\eta$	$k_{\mathrm{ca}}\eta$	ν_1 in original paper
k_{a}	$k_{\mathrm{ca}}\,(1-\eta)$	$k_{\mathrm{ca}}\,(1-\eta)$	ν_2 in original paper
k_{r}	$1/7\,\mathrm{d}^{-1}$	$1/7\,\mathrm{d}^{-1}$	η in original paper
T_{i}	10.5 d	8 d	$T_{\mathrm{i}} = \frac{1}{\mu}$, μ given in original paper
c_1	0.135	0.194	χ_2 in original paper
c_2	1	1	χ_3 in original paper
R_0	2.57	3.39	$\frac{(c_1+k_{\mathrm{ca}})(c_1+k_{\mathrm{e}})(c_1+k_{\mathrm{r}})}{k_{\mathrm{ca}}k_{\mathrm{e}}(c_1+k_{\mathrm{r}}+k_{\mathrm{a}})}\left(1+\frac{k_{\mathrm{a}}}{k_{\mathrm{r}}}\right)$

```
# Specify url:s
url_covid_infected = "https://raw.githubusercontent.com/octonion/COVID
    -19/master/csse_covid_19_data/csse_covid_19_time_series/
    time_series_covid19_confirmed_global.csv"
url_covid_deaths = "https://raw.githubusercontent.com/octonion/COVID-19/
    master/csse_covid_19_data/csse_covid_19_time_series/
    time_series_covid19_deaths_global.csv"
url_covid_recovered = "https://raw.githubusercontent.com/octonion/COVID
    -19/master/csse_covid_19_data/csse_covid_19_time_series/
    time_series_covid19_recovered_global.csv"
#
# Download csv file to current directory
download(url_covid_infected,"covid_confirmed_global.csv")
download(url_covid_deaths,"covid_deaths_global.csv")
download(url_covid_recovered,"covid_recovered_global.csv")
```

The `*.csv` files can then be imported into Julia using packages `CSV, DataFrames`:

```
using CSV, DataFrames
#Import CSV data
data_confirmed = CSV.read("covid_confirmed_global.csv",DataFrame)
data_deaths = CSV.read("covid_deaths_global.csv",DataFrame)
data_recovered = CSV.read("covid_recovered_global.csv",DataFrame)
```

The concept of `DataFrames` originates from computer language R. In Python, this concept is known as `pandas`.

A.2 Model parameters

Table 4 lists model parameters for the model in Problems 2–5.

The infection frequency factor $k_{\mathrm{i}}(t)$ is computed according to Eq. 28 in the original model, where T_1 of Eq. 28 is given in Appendix A.3.

Table 5: Typical operating conditions for reformulation of extended SEIR model in Liu, Magal, Seydi & Webb (2020).

Variable	Value Italy	Value Spain	Comment
t_0	Feb 10, 2020	Feb 3, 2020	First registered infected
S_0	N_{it}	N_{es}	Total population as in Table 4
E_0	$\frac{c_1+k_{\text{ca}}}{k_{\text{e}}}$	$\frac{c_1+k_{\text{ca}}}{k_{\text{e}}}$	Initial estimate for $E(t_0)$
I_0	$\frac{c_1 c_2}{k_{\text{c}}}$	$\frac{c_1 c_2}{k_{\text{c}}}$	Initial estimate for $I(t_0)$
C_0	1	1	Initial estimate for $C(t_0)$
A_0	$\frac{k_{\text{a}}}{c_1+k_{\text{r}}} I_0$	$\frac{k_{\text{a}}}{c_1+k_{\text{r}}} I_0$	Initial estimate for $A(t_0)$
R_0	0	0	Initial estimate for $R(t_0)$
t_1	Mar 16, 2020	Mar 20, 2020	N in original paper

A.3 Operating conditions

Table 5 lists initial values and other operating conditions for the model in Problems 2–5.

References

Keeling, M. J. & Rohani, P. (2008), *Modeling Infectious Diseases in Humans and Animals*, Princeton University Press, 41 William Street, Princeton, New Jersey 08540.

Kermack, W. & McKendrick, A. (1927), 'A contribution to the mathematical theory of epidemics', *Proceedings of the Royal Society A.* **115**(772), 700–721.

Liu, Z., Magal, P., Seydi, O. & Webb, G. (2020), 'A model to predict covid-19 epidemics with applications to south korea, italy and spain', *SIAM News* .
URL: *https://sinews.siam.org/Details-Page/a-model-to-predict-covid-19-epidemics-with-applications-to-south-korea-italy-and-spain*

Liu, Z., Magal, P. & Webb, G. F. (2020), 'Predicting the number of reported and unreported cases for the COVID-19 epidemics in china, south korea, italy, france, germany and united kingdom'.

Martcheva, M. (2015), *An Introduction to Mathematical Epidemiology*, Vol. 61 of *Texts in Applied Mathematics*, Springer, New York.

Sanche, S., Lin, Y. T., Xu, C., Romero-Severson, E., Hengartner, N. & Ke, R. (2020), 'The novel coronavirus, 2019-ncov, is highly contagious and more infectious than initially estimated', *medRxiv* pp. 1–12.

Schwabe, C., Riemann, H. & Franti, C. (1977), *Epidemiology in Veterinary Practice*, Lea & Febiger. pp. 258–260.