Modeling the spread of epidemics and the impact of mobility restrictions

Hendriks, A.H.H. (Teun)

Award date:
2021

EINDHOVEN UNIVERSITY OF TECHNOLOGY

MASTER'S THESIS

MODELING THE SPREAD OF EPIDEMICS

AND THE IMPACT OF MOBILITY RESTRICTIONS

Author:
A.H.H. (Teun)
HENDRIKS
0905393

Supervisors:
M.M. (Martijn)
GÖSGENS MSc.

Prof. dr. R.W. (Remco)
VAN DER HOFSTAD

Prof. dr. P.P.A.M. (Paul)
VAN DER SCHOOT

Department of Applied physics, group of Soft Matter and Biological Physics

Department of Industrial and Applied Mathematics, group of Probability Theory and Statistics

April 6, 2021

Abstract

MODELING THE SPREAD OF EPIDEMICS

by Teun HENDRIKS

The COVID-19 pandemic led to a renewed interest in modeling epidemics with mathematical models across various research fields. We propose an epidemic model based on often-used compartmental models such as the SIR and SEIR models, incorporated with mobility between regions. This results in a so-called remote-force-of-infection epidemic model. We analyze the influence of model parameters on the model dynamics and the spreading velocity in one and two spatial dimensions. We observe traveling epidemic waves with similar properties as solitary waves for one or two spatial dimensions. However, our traveling waves annihilate during a collision. Analytically, we describe the asymptotic behavior of our model, resulting in a modified diffusion equation and a modified telegraph equation. We find analytical expressions describing the propagation velocity of the epidemic wave, which we verify numerically. Next, we investigate how to balance the trade-off between mobility restrictions and transmissions of COVID-19 in a case-study in The Netherlands. We propose a model extension to incorporate so-called super spreaders. We quantify the trade-off between blocking mobility and spreading the infection in terms of a single parameter. Our epidemic model evaluates mobility restrictions' effectiveness where the country is divided into separate regions. We propose heuristic methods to approximate optimal choices for these regions. The results show that our mobility restrictions are highly effective when the infections are concentrated around a few municipalities.

Contents

List of Abbreviations

SIR	Susceptible-Infected-Recovered
SEIR	Susceptible-Exposed-Infected-Recovered
SI	Susceptible-Infected
SIS	Susceptible-Infected-Susceptible
SIRS	Susceptible-Infected-Recovered-Susceptible
SEIRS	Susceptible-Exposed-Infected-Recovered-Susceptible
ODE	Ordinary differential equation
ODEs	Ordinary differential equations
PDE	Partial differential equation
PDEs	Partial differential equations

Chapter 1

Introduction

The COVID-19 pandemic has, by mid-March 2021, caused well over 100 million infected individuals worldwide, of which more than 2 million passed away in over 200 countries. Many countries across the globe are taking measures to contain the virus, such as social-distancing, national lock-downs, curfews, and more.

Similar to other infectious diseases, such as the H1N1 outbreak in 2009 [1], the SARS epidemic in 2003 [2], and influenza epidemics [3, 4], the geographical spread is determined by human behavior and human mobility patterns across multiple spatial scales [5]. While infectious contacts of individuals give rise to local outbreaks of a disease, the combination with mobility of people causes spatial spread [6].

To understand, for example, how fast diseases spread [7], find the outbreak location [8], or which regions would benefit from mobility restrictions to contain an outbreak [9], researchers use mathematical models to describe epidemics spreading through a population.

Single population epidemic models [10, 11] divide a population into different compartments. For example, the so-called SIR model divides individuals into *susceptible*, *infected*, and *recovered* compartments, see Figure 1.2. Another example is the *susceptible-exposed-infected-recovered* (SEIR) model.

In many diseases, there is a significant period between getting infected and actually being infectious. This period is called the latent period. In this period, people are considered exposed. Example diseases with such a latent period are Ebola [12], measles [13], and COVID-19 [14].

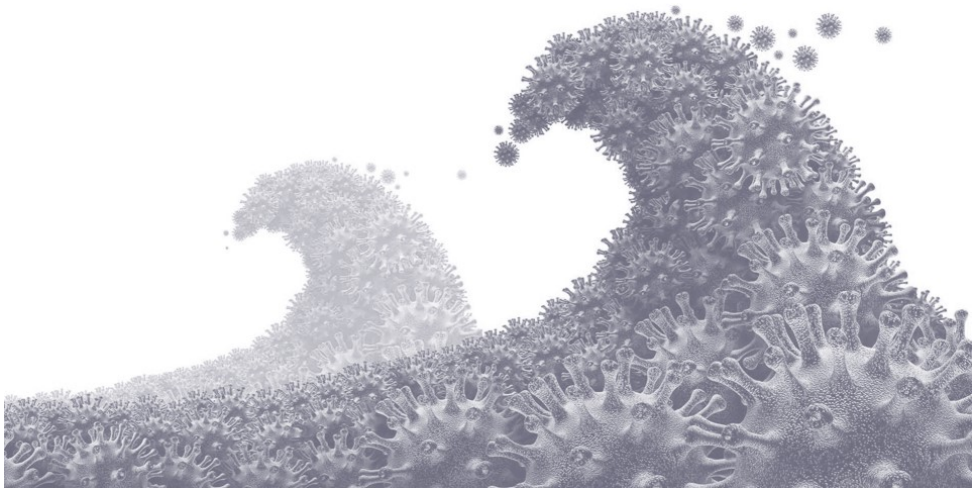


FIGURE 1.1: COVID-19 waves. Image provided by *Getty Images* [15].

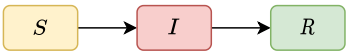


FIGURE 1.2: Susceptible-infected-recovered compartment model. Individuals are initially susceptible to the disease. They can get infected through contact with infectious individuals. After an infectious period, individuals become removed, which means they have recovered or have passed away.

While these classical compartmental models can describe the disease's exponential growth, they do not capture spatial effects. A class of epidemic models that includes spatial differences is based on reaction-diffusion models [16, 17]. The combination of diffusive dispersal (mobility process) *between* different populations and local exponential growth *within* a single population resembles observed epidemic behavior.

A closely related class of models is based on the remote-force-of-infection. These models have the advantage that the mobility process does not need to be modeled explicitly [18, 19, 20]. Instead, only the mobility process consequences, i.e., new infections, are modeled rather than individuals moving to different locations.

In these models, the population is subdivided into multiple connected subpopulations representing, for example, regions of a country. The connection between regions resembles the mobility between them. A network of subpopulations is called a metapopulation. Models using the metapopulation framework are usually very coarse. Some more detailed models have been used to describe epidemics [21, 22, 23].

These detailed so-called agent-based models try to capture and model individuals referred to as agents. Agents can be part of households, local communities, schools, and/or work-groups. A disadvantage of these detailed models is that they use many assumptions and parameters, making the analysis of the model and its parameters harder. In general, these detailed models are better suited for small-scale simulations as they require a lot more computational power In contrast, the metapopulation compartment models are often used for large-scale analysis [4].

In this work, we focus on large-scale analysis of epidemics that spreads through populations. This work is motivated by the current COVID-19 pandemic. We analyze and apply epidemic models to evaluate the effectiveness of mobility restrictions on a single nation's scale. For this, we want to understand the dynamics of the epidemic model and investigate the influence of different model parameters. This can be summarized in the following research questions:

1. How is the dynamic of an epidemic model and specifically the spreading velocity influenced by model parameters?
2. How to balance the trade-off between mobility restrictions and transmission of COVID-19?

We consider restrictions where we divide a country into different regions. We block the mobility between different regions but allow for mobility within these regions. This raises two sub-questions for research question 2:

(a) How large should a 'region' be for effective containment of the disease?

(b) How to evaluate the effectiveness of mobility restrictions?

1.1 Methodology

This work focuses on epidemic models using a remote-force-of-infection and different metapopulation networks. The actual movement of individuals is not modeled, but the results of these movements, i.e., the disease's transmission, are.

We consider multiple compartment arrangements, but we focus mainly on SIR and SEIR models. The SEIR model is used to model COVID-19, and it is closely related to the SIR model. Other compartment setups are discussed briefly.

Stochastic effects in such models play a role when the systems are small, but when the number of individuals is large, the stochastic effects can be neglected. We mainly focus on large systems, where stochastic effects are negligible, and a deterministic mean-field approach describes the epidemic well.

In real-world applications, the underlying mobility networks can be complex. We can observe wave-like patterns for relatively simple networks, such as a one-dimensional network or a two-dimensional grid.

We analyze these patterns both numerically and analytically for a one-dimensional metapopulation network. In particular, the focus is on the wave profile, wave velocity, and outbreak threshold. The threshold determines whether the disease dies out early or grows exponentially in size. We verify analytical results numerically.

Shortly after a concentrated outbreak, the wave profile in an SIR model can be approximated by a modified diffusion equation. In the case of an SEIR model, by a modified telegraph equation [24]. The telegraph equation also arises when studying the propagation of electrical signals in a cable or transmission line, but it appears in more branches of science [25, 26].

After the diffusive regime, a traveling wave with constant shape and propagation velocity appears. In the literature, waves with similar properties are called solitary waves [27]. We show that the traveling waves we observe are different from solitary waves.

Next, we compare these findings with results obtained on a two-dimensional grid. We compare wave-like patterns and the speed at which the disease spreads. Even for two initial outbreak locations, we show that the wavefront eventually becomes circular.

Finally, as an application of our model, we evaluate strategies that impose mobility restrictions upon subpopulations. We use The Netherlands as an example. Mobility is estimated from a dataset, which contains the average number of traveling people between Dutch municipalities.

The number of new infections is weighted against loss of mobility for different mobility restriction strategies. As mobility goes hand in hand with economic activity, measures that restrict mobility have a potentially large social-economic impact [28].

We quantify the trade-off between the spreading of the disease and mobility in a single parameter. Based on this parameter, we present mobility restriction strategies to minimize the number of new infections while maximizing mobility. We propose a method for determining the mobility restrictions, which is inspired by community detection.

We analyze stochastic effects in the spreading of the epidemic. For COVID-19, super-spreaders play a role in spreading the disease [29, 30, 31]. They transmit the disease to many others. We propose an epidemic model, which incorporates super-spreading effects.

All the code to reproduce the experiments is publicly available at `https://github.com/teunhend/master-thesis`.

1.2 Thesis outline

Chapter 2 provides a theoretical background and an overview of different epidemic models. In Chapter 3 we present analytical results regarding the models. We provide analytical solutions that are valid shortly after an outbreak. Also, we present expressions for the wave propagation velocity. In Chapter 4, we verify our analytical results numerically for a one-dimensional network. We show results on a two-dimensional network. Chapter 5 presents the results related to our second research question. We introduce the mobility patterns of The Netherlands. We select mobility restrictions and evaluate these restrictions. The last chapter provides a summary of conclusions and suggestions for future work.

Chapter 2

Theory and background

This chapter serves as a theoretical background for this work. The most well-known classical epidemic model is developed by Kermack and McKendrick and published in 1927 [10]. As this article is fundamental to the field, it is republished in 1991 [32].

The idea behind their model is to divide a population into different compartments: *Susceptible*, *Infected*, and *Removed*. The model is referred to as the SIR-model; see Figure 1.2. This setup allows for the description of infectious diseases where individuals are resistant to the disease after recovery.

This is not the case for every disease. For example, individuals do not build up a life-long resistance to the common cold [33]. By varying the compartment setup, we can apply the ideas of the SIR model to different kinds of infectious diseases.

In Section 2.1.1, we describe some possible variations of the SIR model and the applications of these variations. In Section 2.1.2, we explain the SIR model in more depth, as it is a good introduction to field of modeling epidemics. In Section 2.1.3, we explain a compartmental model relevant for COVID-19 in more detail. We discuss model behavior in Section 2.1.4.

In Section 2.2.1, we describe how to include a spatial component in the classical epidemic models. Often, reaction-diffusion models are used, which we describe in Section 2.2.2. For The Netherlands, commuters dominate the daily mobility patterns. For these patterns, we show that reaction-diffusion models reduce to so-called remote-force-of-infection models in Section 2.2.3.

Lastly, in Section 2.3, we discuss the stochastic effects for epidemics.

2.1 Classic epidemic models

In this section, we show different variations and applications of epidemic models similar to the classical SIR model. We discuss the SIR model as an introduction to the field of epidemic models. We analyze its behavior. Lastly, we discuss the *Susceptible Exposed Infected Recovered* (SEIR) model in more detail, as this model is relevant to COVID-19.

2.1.1 Model variations

Extensions and variations are developed to describe different infectious diseases. For example, by modifying what happens after recovery of the disease or by adding a fourth compartment: *Exposed*. We show some variations in Table 2.1 and discuss the different models in the remainder of this section.

We can choose to include a latent period to model the period between getting infected and becoming infectious. This state is modeled by an *Exposed* period. An example model is the Susceptible-Exposed-Infected-Recovered (SEIR) model. See Figure 1.2 versus Figure 2.4.

We consider three different outcomes after recovery of a disease:

1. Individuals have a permanent resistance to the disease after recovery. Examples are the SIR and SEIR models, see Figure 1.2 and Figure 2.4.
2. Individuals have built up no resistance after recovery. Example models are the Susceptible Infected (SI) model, where people remain infectious and never recover, and the Susceptible Infected Susceptible (SIS) model, see Figure 2.1 and Figure 2.2.
3. Individuals lose resistance over time. An example model is the SEIRS model. This model is similar to the SEIR model, but individuals become susceptible again after some period in which they lose resistance to the disease, see Figure 2.5.

An epidemic model should match a disease, but also the time scale of interest. For example, in Table 2.1, influenza is modeled using an SIR model and an SIRS model. In the first case, influenza is modeled in a single year. Within this year, individuals do not lose their resistance to influenza [34]. But when multiple years are considered, an SIRS model is more suited, as people are susceptible to next year's variant of influenza [35].

Model name	Lasting immunity	Compartment diagram	Example applications
SIR	yes	Figure 1.2	Modeling influenza [34]
SI	no	Figure 2.1	Modeling HIV [36]
SIS	no	Figure 2.2	Modeling tuberculosis [37]
SIRS	temporal	Figure 2.3	Modeling influenza [35]
SEIR	yes	Figure 2.4	Modeling COVID-10 [9]
SEIRS	temporal	Figure 2.5	Modeling computer malware [38]

TABLE 2.1: Overview of different compartmental models and applications.

In the case of COVID-19, it remains unclear if and for how long immunity lasts after recovery of the disease [39, 40]. However, there is a latent period present for COVID-19 infections [41]. In Chapter 5, we consider a total time of 30 days, in which it is expected that individuals do not lose their resistance to COVID-19 [39]. Therefore we focus on the SEIR model for COVID-19 related results.

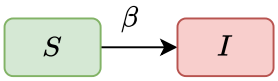


FIGURE 2.1: Schematic overview of a Susceptible-Infectious model. After transmission of the disease, individuals become and remain infectious.

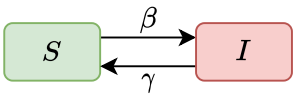


FIGURE 2.2: Schematic overview of a Susceptible-Infectious-Susceptible model. After the infectious period, individuals become susceptible again.

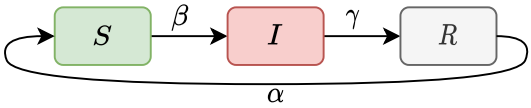


FIGURE 2.3: Schematic overview of a Susceptible-Infectious-Removed-Susceptible model. After the infectious period, individuals have a temporal resistance to the disease.

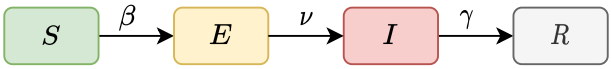


FIGURE 2.4: Schematic overview of a Susceptible-Exposed-Infectious-Removed model. In this case, after getting infected, individuals are not directly contagious to others. The are exposed and become eventually infectious with after a latent period $1/\nu$.

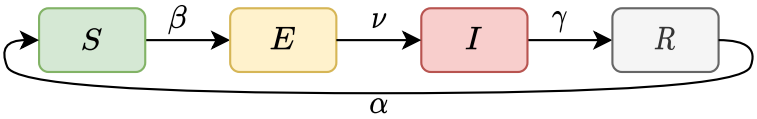


FIGURE 2.5: Schematic overview of a Susceptible-Exposed-Infectious-Removed-Susceptible model. This model is similar to the SEIR model in Figure 2.4, but individuals lose their resistance to the disease with rate α and become susceptible again.

Important parameters which we use in the remainder of this work are the average periods individuals spend in the exposed and infected compartments. We call these periods the *latent period*, $1/\nu$, and the *infectious period*, $1/\gamma$, respectively.

As mentioned before, we discuss the model equations of the SIR and SEIR models in the next sections. For the other models, we refer to Appendix A. In Section 2.1.4, we briefly compare the behavior of the models from Table 2.1.

2.1.2 SIR epidemic model

The first compartment in the SIR model is the group of healthy individuals. While these individuals are currently healthy, they are prone to contract the disease and are therefore called *susceptible*. We denote the number of susceptible individuals by S.

The second compartment consists of individuals who currently have the disease and are contagious to other people. They are called *infectious*. This group allows the disease to spread. Contact between a contagious individual and a susceptible individual may lead to the infection of the susceptible. We denote the number of infectious individuals by I.

Lastly, a group of individuals have had the disease and are now resistant to the disease. These individuals are called *recovered*. We denote the number of removed individuals by R.

The number of individuals in the different compartments changes over time. So $S(t)$, $I(t)$, and $R(t)$ are all functions of time t. We assume the total population size N to be constant over time, i.e., individuals cannot leave or enter the system. N can be calculated by summing the different compartments:

$$N = S(t) + I(t) + R(t).$$

Kermack and McKendrick propose ordinary differential equations (ODEs) that describe how the number of individuals in each compartment change over time. Transmission of the disease can only happen when a susceptible individual gets in contact with an infectious individual. During this contact, the disease is transmitted to the susceptible person with a certain probability.

We use *incidence* to refer to how many people move from susceptible to being infectious per unit of time. Incidence depends on the transmission probability and on the number of contacts people have per unit of time. We assume the following:

- The number of contacts per unit of time an individual has is equal to C. We assume this is independent of the population size; this is reasonable for larger populations.
- The probability that a given contact of a susceptible individual is with an infectious individual is $\frac{I}{N}$, also known as *perfect mixing*. So, the expected number of connections with infectious individuals per susceptible individual per unit of time is equal to $C\frac{I}{N}$.
- The transmission probability of the disease is denoted by ε. It is equal to the probability of the susceptible individual getting infected with the disease during contact with an infectious individual. Then, the rate at which one susceptible person gets infected is then equal to $\varepsilon C\frac{I}{N}$.
- The incidence is equal to the number of susceptible individuals times the rate described above, which equals $\varepsilon C\frac{SI}{N}$. The infection rate β is equal to εC. We can thus write the incidence as $\beta\frac{SI}{N}$.

Since we assume that the number of contacts a person has is independent of the population size, the type of incidence is called *normal incidence*. Next, we define the force-of-infection λ as $\lambda = \beta\frac{I}{N}$.

In Section 2.2.3, models based on a remote-force-of-infection are described. In these models, multiple populations exert a remote force of infection upon each other, which results in the transmission of the disease to different populations.

As incidence is defined as how many people move from being susceptible to being infectious per unit of time, we find the derivative of the number of susceptible individuals with respect to time to be equal to minus incidence

$$\frac{dS}{dt} = -\beta \frac{SI}{N}. \tag{2.1}$$

These people move into the infectious compartment. People leave the infectious compartment when they recover or pass away. Kermack and McKendrick assume this to be a constant per capita rate per unit of time.

Let γ be the recovery rate. The inverse of the recovery rate is equal to the infectious period. Combining influx and outflux of contagious individuals, we find

$$\frac{dI}{dt} = \beta \frac{SI}{N} - \gamma I. \tag{2.2}$$

Finally, infectious individuals eventually move to the removed compartment,

$$\frac{dR}{dt} = \gamma I. \tag{2.3}$$

The equations above combine into a system of ODEs:

$$\frac{dS}{dt} = -\beta \frac{SI}{N}, \tag{2.4}$$
$$\frac{dI}{dt} = \beta \frac{SI}{N} - \gamma I, \tag{2.5}$$
$$\frac{dR}{dt} = \gamma I. \tag{2.6}$$

This system comes with initial conditions $S(t_0)$, $I(t_0)$, and $R(t_0)$. Note that we did not write down the time dependence of S, I, and R in all equations explicitly. There exist analytical solutions for this system of ODEs [42].

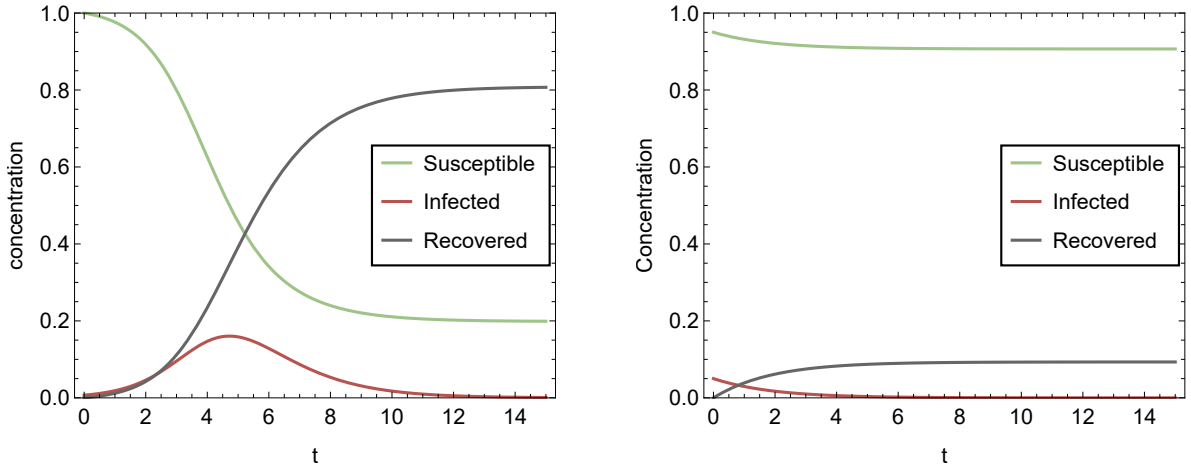
We can simplify the system of ODEs if we scale it with the population size N and define a dimensionless time $t^* = \gamma t$. After dropping the asterisk, we obtain the following equations:

$$\frac{ds}{dt} = -\mathcal{R}_0 si, \tag{2.7}$$
$$\frac{di}{dt} = (\mathcal{R}_0 s - 1)i, \tag{2.8}$$
$$\frac{dr}{dt} = i, \tag{2.9}$$

where $s = \frac{S}{N}$, $i = \frac{I}{N}$, and $r = \frac{R}{N}$. Equation (2.9) can be replaced by the conservation equation: $r = 1 - s - i$. The only parameter in this system of ODEs is $\mathcal{R}_0 = \frac{\beta}{\gamma}$. This systems requires initial conditions $s(t = 0)$, $i(t = 0)$, and $r(t = 0)$.

$\mathcal{R}_0$ is the so-called *Basic Reproduction Number*. $\mathcal{R}_0$ and is important for the model behavior. In Figure 2.6, we present two results for different values of $\mathcal{R}_0$. Before we discuss the model

behavior in more detail the Section 2.1.4, we introduce the model equations for the Susceptible-Exposed-Infected-Recovered (SEIR) model in the next section.



(a) $\mathcal{R}_0 = 2$ (b) $\mathcal{R}_0 = 0.5$

FIGURE 2.6: Solutions of (2.7), (2.8), and $r = 1 - s - i$ for different values of the model parameter $\mathcal{R}_0$. Both results are based on initial conditions $s(t = 0) = 0.95$, $i(t = 0) = 0.05$, and $r(t = 0) = 0$.

2.1.3 SEIR epidemic model

The Susceptible-Exposed-Infectious-Susceptible model has four compartments. Once infected, individuals are not immediately infectious but instead are exposed. After a latent period, individuals become eventually infectious.

In the SEIR model, we denote the rate at which people move from the exposed to the infectious compartment by ν. This means that the average duration of the latent period, the time spent in the exposed compartment, is equal to $1/\nu$. We follow similar steps as described in the previous section to obtain the following equations:

$$\frac{ds}{dt} = -\beta si, \tag{2.10}$$

$$\frac{de}{dt} = \beta si - \nu e, \tag{2.11}$$

$$\frac{di}{dt} = \nu e - \gamma i, \tag{2.12}$$

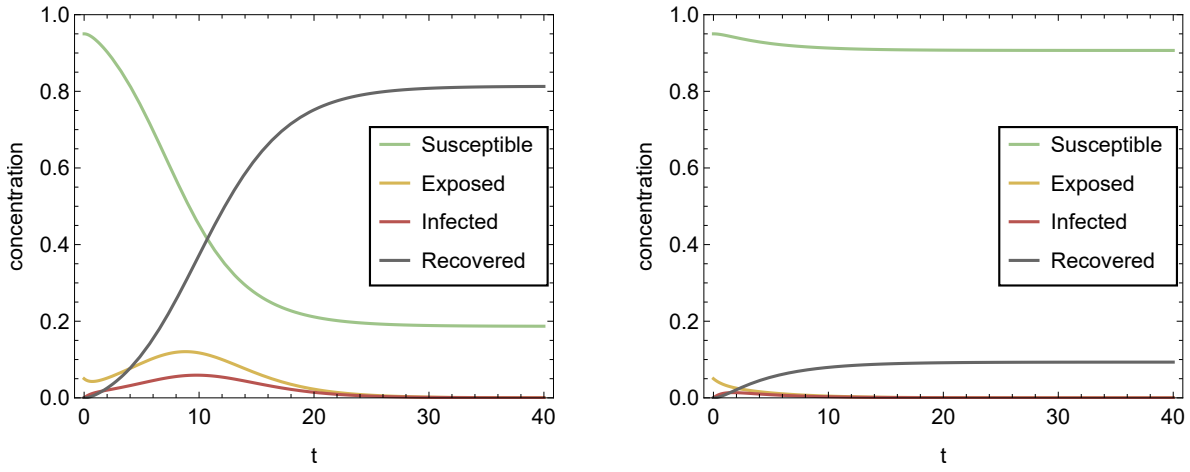
$$\frac{dr}{dt} = \gamma i. \tag{2.13}$$

We define a dimensionless time $t^* = \gamma t$. When we substitute this dimensionless time and after dropping the asterisk, we obtain the following equations:

$$\frac{ds}{dt} = -\mathcal{R}_0 si, \tag{2.14}$$
$$\frac{de}{dt} = \mathcal{R}_0 si - ae, \tag{2.15}$$
$$\frac{di}{dt} = ae - i, \tag{2.16}$$
$$\frac{dr}{dt} = i, \tag{2.17}$$

where $\mathcal{R}_0 = \beta/\gamma$ and $a = \nu/\gamma$. The basic reproduction number of an SEIR model is equal to the $\mathcal{R}_0$ in an SIR model for equal β and γ. The parameter a can be interpreted as the ratio of the infectious period over the latent period. For large values of a, the SEIR model can be approximated by the SIR model.

The systems has initial conditions $s(t = 0)$, $e(t = 0)$, $i(t = 0)$, $r(t = 0)$ and can be solved analytically [43]. The model behavior is similar to the SIR model, as we will discuss in the next section.



(a) $\mathcal{R}_0 = 2$ (b) $\mathcal{R}_0 = 0.5$

FIGURE 2.7: Solutions of (2.14), (2.15), (2.16), and $r = 1 - s - e - i$, where $a = 1$ and for different values of the model parameter $\mathcal{R}_0$ and . Both results are based on initial conditions $s(t = 0) = 0.95$, $e(t = 0) = 0.05$, $i(t = 0) = 0$, and $r(t = 0) = 0$.

2.1.4 Model behavior

In this section, we explain the influence of the *Basic Reproduction Number*, $\mathcal{R}_0$, on the dynamics of the epidemic. We introduce the reproduction number communicated by many governments and health authorities to the general public as the *Effective Reproduction Number*, $\mathcal{R}_e(t)$ [44]. We also discuss at the mathematical properties of the model and the so-called *final size* of the epidemic. These aid us in analyzing the results in this work. Lastly, we show different model behavior for models presented in Table 2.1.

The threshold in an epidemic model is based on the value of the basic reproduction number, $\mathcal{R}_0$. It is defined as the number of expected successful transmissions of the disease in a completely susceptible population by a singly infectious individual [45]; see Figure 2.8.

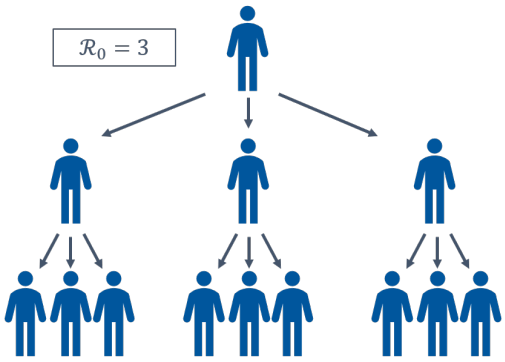


FIGURE 2.8: Visualization of a basic reproduction number $\mathcal{R}_0 = 3$. Infectious individuals each transmit the disease to three susceptible individuals. Hereby, on average, they cause three new cases.

From (2.8), we can see that there can only be exponential growth in the number of infectious individuals if $\mathcal{R}_0 > 1$. Note that $0 \leq s \leq 1$. However, from 2.7(a), we also see that exponential growth eventually turns into exponential decay.

From(2.8) we also see that there is only actual exponential growth when $s(t)\mathcal{R}_0 > 1$. We define the effective reproduction number $\mathcal{R}_e(t)$ as $\mathcal{R}_e(t) = s(t)\mathcal{R}_0$. There is exponential growth at time t during an epidemic outbreak if $\mathcal{R}_e(t) > 1$, while when $\mathcal{R}_e(t)$ is less than 1, there is an exponential decay at time t. We can also see this behavior in Figure 2.6.

From this, we can derive the notion of *herd immunity*. This is defined as the fraction of the population that needs to be resistant to the disease such that there is no exponential growth in the number of infectious individuals possible anymore. It is only defined for $\mathcal{R}_0 > 1$. We have herd immunity when $\mathcal{R}_e \leq 1$. In other words, $s\mathcal{R}_0 \leq 1$. This means that $s \leq \frac{1}{\mathcal{R}_0}$. Note that $1 - r \leq s$, which results in the following equation for the fraction of recovered individuals r:

$$r \geq 1 - \frac{1}{\mathcal{R}_0}. \tag{2.18}$$

For example, for COVID-19, $\mathcal{R}_0$ is estimated at 2.5 [46]. This means that herd immunity is achieved when $r = 0.6$. In other words, when 60% of the people are resistant.

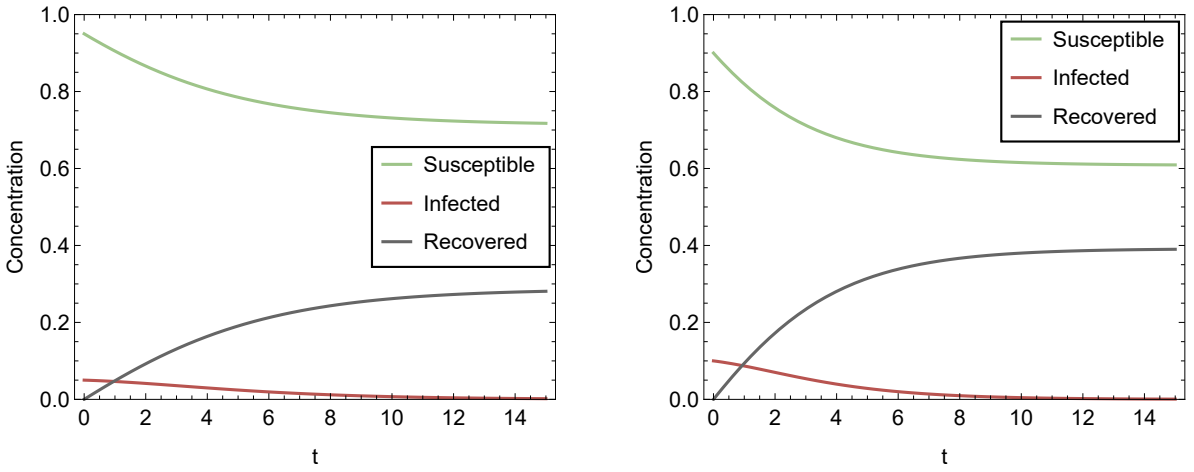
In Figure 2.9, we see two solutions of (2.7), (2.8), and $r = 1 - s - i$ with different initial conditions for $\mathcal{R}_0 = 1$. While the infectious concentration $i(t)$ starts to decay slowly, the fraction of the population that ends up in the recovered compartment is different. The fraction $r(t)$ for t when the epidemic has died out is called the final size of the epidemic. Figure 2.9 shows us that the final size not only depends on the model parameter $\mathcal{R}_0$, but also on the initial conditions.

As the derivative in (2.7) is always negative, $s(t)$ is a monotonically decreasing function. For $s(t)$ to have a physical meaning, we have $0 \leq s \leq 1$. These properties combined results in

$$\lim_{t \to \infty} s(t) = s_\infty. \tag{2.19}$$

Here, s_∞ is the fraction of individuals in the susceptible compartment after an epidemic outbreak. Similarly, we can see in Figures 2.6 and 2.9 that

$$\lim_{t \to \infty} i(t) = 0. \tag{2.20}$$



(a) $i(t=0)=0.05$ (b) $i(t=0)=0.1$

FIGURE 2.9: Solutions of (2.7), (2.8), and $r = 1 - s - i$ for $\mathcal{R}_0 = 1$. (a): Result is based on initial conditions $s(t=0) = 0.95$, $i(t=0) = 0.05$, and $r(t=0) = 0$. (b): Result is based on initial conditions $s(t=0) = 0.9$, $i(t=0) = 0.1$, and $r(t=0) = 0$.

Using the conservation equation $r = 1 - s - i$, we find

$$\lim_{t\to\infty} r(t) = 1 - s_\infty. \tag{2.21}$$

For the susceptible-infected-removed (SIR) model, the final size equation is given by :

$$r_\infty = 1 - s(t=0)\exp\Big[-\mathcal{R}_0 r_\infty\Big]. \tag{2.22}$$

From this equation, we can see when $s(t=0)$ becomes smaller, or equivalently, $i(t=0)$ is larger, solutions r_∞ become larger. This is in line with Figure 2.9. For the derivation of (2.22), we refer to Appendix B.

Another interesting value we will use in our results analysis is the peak concentration of infectious individuals. This peak number is defined as

$$i_{\text{peak}} = \max_t i(t). \tag{2.23}$$

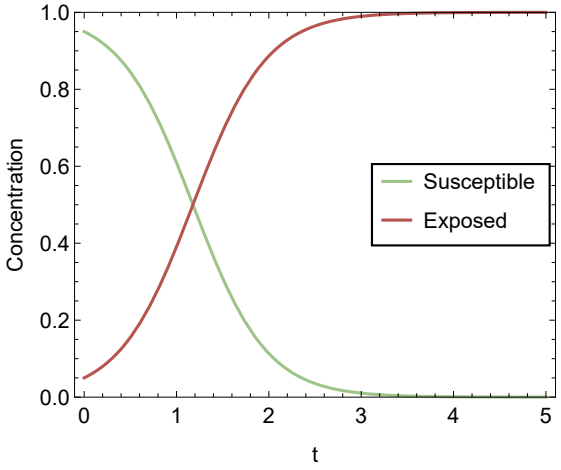
In Figures 2.6 and 2.9, the peak concentration is the maximum value obtained by the red curve. The peak concentration tells something about the epidemic's intensity.

The final size of the epidemic in the SEIR model can also be calculated via (2.22). However, the peak concentration (2.23) is lower as a fraction of the infected individuals is only exposed and not yet infectious. This lower concentration of infectious individuals is sustained for longer periods of time.

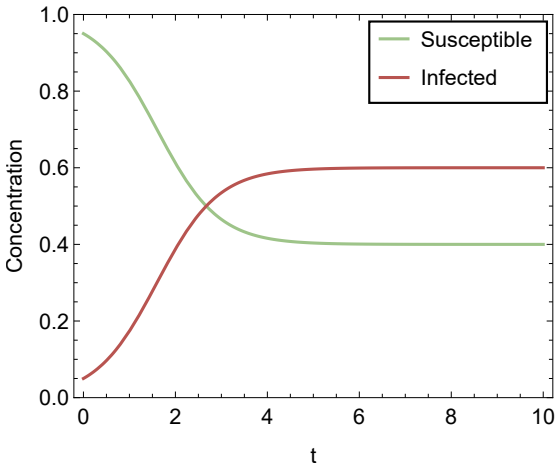
Figure 2.6(a) and Figure 2.7(a) clearly show this. When the epidemic has died out, in other words, $i(t) = 0$, the value of $r(t)$ is in both models at approximately 0.8. While the peak intensity of $i(t)$ in Figure 2.7(a) is lower, it is sustained for longer periods of time compared to Figure 2.6(a).

The SIR and SEIR models both result in life-long resistance to a disease. In the SI model, individuals become infectious and never recover. This eventually results in everybody becoming

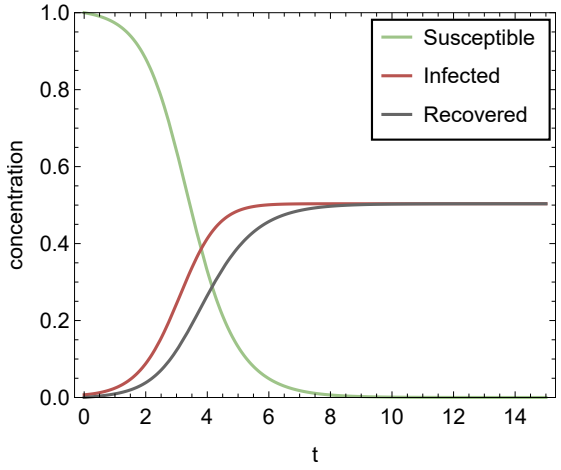
infectious, see Figure 2.10(a). In the SIS model, individuals have no resistance to the disease after recovery. In the SIRS and the SEIRS model, lose the resistance to the disease after recovery. This results in a sustained fraction of people that are infectious, see Figure 2.10(b), Figure 2.10(c), and Figure 2.10(d). In these cases, the infectious disease has become endemic and remains active in the population.



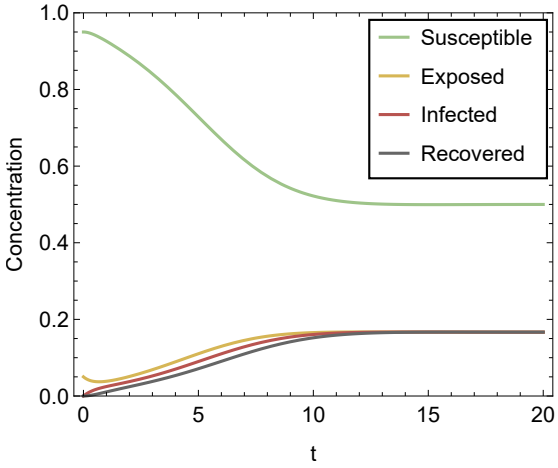
(a) SI model



(b) SIS model



(c) SIRS model



(d) SEIRS model

FIGURE 2.10: Solutions of different models described in Appendix A. (a): Initial conditions are $s(t = 0) = 0.95$ and $i(t = 0) = 0.05$. (b): $\mathcal{R}_0 = 2$ and initial conditions are $s(t = 0) = 0.95$ and $i(t = 0) = 0.05$. (c): $\mathcal{R}_0 = 2$, $\alpha = 1$ and initial conditions are $s(t = 0) = 0.95$, $i(t = 0) = 0.05$, and $r(t = 0) = 0$. Result is based on initial conditions $s(t = 0) = 0.95$, $i(t = 0) = 0.05$, and $r(t = 0) = 0$. (d): $\mathcal{R}_0 = 2$, $\alpha = 1$, and ν=1, and initial conditions are $s(t = 0) = 0.95$, $i(t = 0) = 0.05$, and $r(t = 0) = 0$. Result is based on initial conditions $s(t = 0) = 0.95$, $e(t = 0) = 0.05$, $i(t = 0) = 0$, and $r(t = 0) = 0$.

2.2 Metapopulation epidemic models

In the previous section, we mentioned that we focus on SIR and SEIR models. In our results, we use the Dutch mobility data in combination with our models. In The Netherlands, traveling individuals are typically commuters who travel between a destination and their base location.

In this chapter, we show that a remote-force-of-infection model suits these properties well. Firstly, we explain the metapopulation framework in Section 2.2.1. Then, we discuss often used reaction-diffusion models in Section 2.2.2 and show how these models reduce to remote-force-of-infection models when we consider mobility patterns of commuters. We discuss remote-force-of-infection models in Section 2.2.3.

2.2.1 Metapopulation framework

The spatial population distribution and the interplay between mobility patterns and the mixing of people cannot be captured by models described in Section 2.1. Those models consider a single population with no spatial differences within its population. With a model formulated in the framework of a so-called *meta-population*, it is possible to introduce spatial differences.

We can divide a single population into many different subpopulations. For example, based on gender, age, or location. The subpopulation sizes do not have to be equal. These subpopulations may or may not be interconnected to other subpopulations in other regions. The network of all subpopulations is called a metapopulation [47]. In this work, we divide a population into subpopulations based on their location.

Let $\mathcal{A}$ be the set of all different subpopulations. We assign to each location a number between 1 and n, which is equal to the total number of different locations. Each location is represented by a subpopulation. x is said to be a location if $x \in [1, ..., n] = \mathcal{A}$.

The subpopulation size of location x is equal to N_x. In the case of compartments, we can express the number of people in, for example, compartment S at location x at time t via $S_x(t)$. When considering concentrations, we write for location x: $s_x(t) = S_x(t)/N_x$. Similarly, we have for other compartments $e_x(t) = E_x(t)/N_x$, $i_x(t) = I_x(t)/N_x$, and $r_x(t) = R_x(t)/N_x$.

Mobility networks can be used to set up a metapopulation. In this work, we consider the mobility network of The Netherlands; see Section 5.2.3. Worldwide air travel is another example of a mobility network [48]. In the latter, the relevant time scales are in the order weeks and are in general longer than the relevant timescales within the mobility network of The Netherlands, which are of time scales less than a day.

Nodes in these networks represent different locations with subpopulation sizes N_x. The edges between the nodes represent the connection between to locations. The edge weight tells us how strong this connection is. In other words, the edge weight represents how many people travel between the two subpopulations.

These networks can be highly heterogeneous in subpopulation sizes, number of edges and, edge weights [49]. We show this heterogeneity in the Dutch mobility dataset in Section 5.2.3.

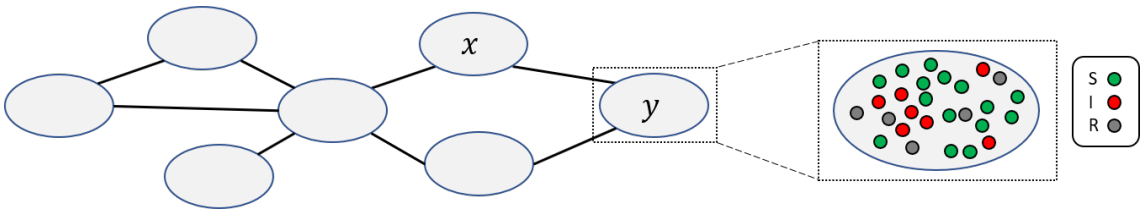


FIGURE 2.11: Example meta-population SIR model, see Figure 1.2. Each subpopulation consists of many individuals who are in one of the four compartments. S: Susceptible, Infected, R: Recovered. Ovals represent different populations. Black lines represent connections between these populations.

2.2.2 Reaction-diffusion model

Epidemic models based on reaction-diffusion models are often used in combination with a metapopulation structure. In these models, we model the spreading of individuals through different subpopulations. We show in this section that reaction-diffusion models can reduce to remote-force-of-infections models. In the latter, we model the *spreading of information*, in other words, the spreading of new infectious cases, while individuals remain fixed to their base location.

Figure 2.11 shows an example of an the SIR model in a metapopulation framework. We extend (2.4), (2.5), and (2.6) with diffusive dispersal among different subpopulations. Dispersal is defined by mobility rates $\omega_{yx} > 0$ from subpopulation x to y; see Figure 2.12. Note that ω_{xy} does not need to be equal to ω_{yx}.

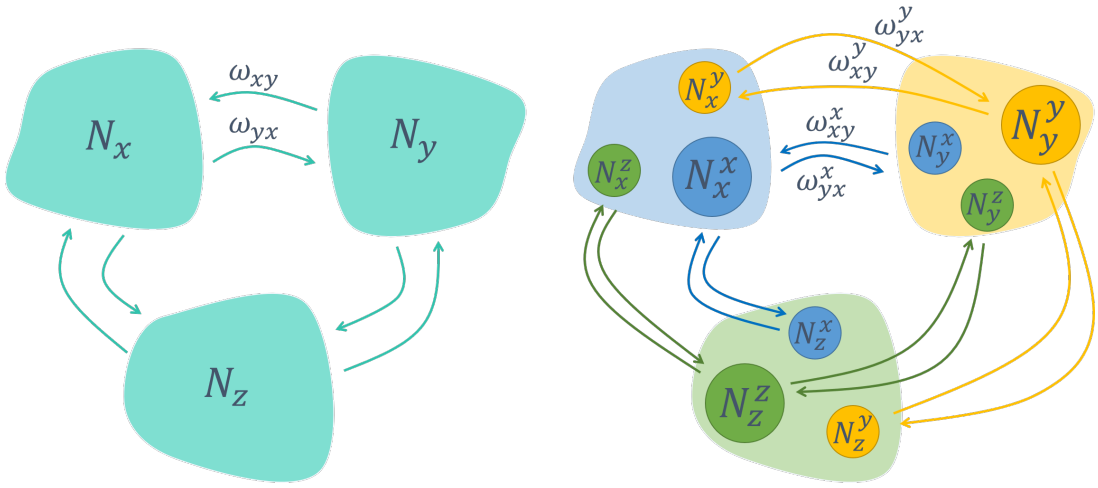


FIGURE 2.12: Meta-population model with mobility. Shapes represent different populations, and arrows represent mobility rates. Left: Diffusive dispersal of indistinguishable individuals between different locations is captured by different mobility rates ω_{xy}. Right: Individual mobility patterns of commuters are captured in the dispersal process. Individuals from base location x travel from and to connected location z with mobility rates ω_{xz}^x and ω_{zx}^x, respectively.

These rates are usually based on mobility data (see Section 5.2.3) or simulated using gravity-like models [50, 51]. The idea behind these models is that the mobility rate between subpopulation x and w is proportional to

$$\omega_{xy} \sim \frac{N_x N_y}{d(x,y)}, \tag{2.24}$$

where $d(x,y)$ is a distance between subpopulation x and y. If these mobility rates are included in (2.4), (2.5), and (2.6), the result is a *meta-population reaction-diffusion* system:

$$\frac{dS_x(t)}{dt} = -\beta \frac{S_x(t)I_x(t)}{N_x} + \sum_{y\in\mathcal{A}} \Big(\omega_{xy} S_y(t) - \omega_{yx} S_x(t) \Big), \tag{2.25}$$

$$\frac{dI_x(t)}{dt} = \beta \frac{S_x(t)I_x(t)}{N_x} - \gamma I_x(t) + \sum_{y\in\mathcal{A}} \Big(\omega_{xy} I_y(t) - \omega_{yx} I_x(t) \Big), \tag{2.26}$$

$$\frac{dR_x(t)}{dt} = \gamma I_x(t) + \sum_{y\in\mathcal{A}} \Big(\omega_{xy} R_y(t) - \omega_{yx} R_x(t) \Big), \tag{2.27}$$

where $\overline{N_x}$ is the population size present location x in the steady-state. $\overline{N_x}$ can be calculated using the following conservation law and steady-state conditions:

$$N_{tot} = \sum_{x \in \mathcal{A}} \Big(S_x(t) + I_x(t) + R_x(t) \Big) \tag{2.28}$$

$$\frac{\overline{N_x}}{\overline{N_y}} = \frac{\omega_{xy}}{\omega_{yx}} \qquad \text{for } x \neq y \text{ and } x, y \in \mathcal{A} \tag{2.29}$$

Lastly, the system has initial condition $S(x, t = 0)$, $I(x, t = 0)$, and $R(x, t = 0)$ for $x \in \mathcal{A}$.

The left side of Figure 2.12 shows the diffusion process. Now we include return rates by keeping track of the individual's base location. This allows us to use commuters' mobility patterns. The right side of Figure 2.12 shows these mobility patterns.

We define the mobility rate ω^z_{xy} as the rate of individuals that travel to x, who are currently in y, and have their base location at z. From now on, we only consider commuters. Hence we assume $\omega^z_{xy} = 0$ when $x \neq z$ or $y \neq z$. This results in traveling individuals between certain destinations and their base location.

Let $S^z_x(t)$, $I^z_x(t)$, and $R^z_x(t)$ be the number of susceptible, infectious, and recovered individuals with base location z who are currently at location x at time t, respectively. For each base location $z \in \mathcal{A}$ we can rewrite (2.25), (2.26), and (2.27). For example, susceptible individuals from z who are currently at x can get infected by infectious individuals from all other $y \in \mathcal{A}$, who are also currently at x, see 2.30. This results in the following equations:

$$\frac{dS^z_x(t)}{dt} = -\beta \frac{S^z_x(t)}{N_x} \sum_{y \in \mathcal{A}} I^y_x(t) + \sum_{y \in \mathcal{A}} \Big(\omega^z_{xy} S^z_y(t) - \omega^z_{yx} S^z_x(t) \Big), \tag{2.30}$$

$$\frac{dI^z_x(t)}{dt} = \beta \frac{S^z_x(t)}{N_x} \sum_{y \in \mathcal{A}} I^y_x(t) - \gamma I^z_x(t) + \sum_{y \in \mathcal{A}} \Big(\omega^z_{xy} I^z_y(t) - \omega^z_{yx} I^z_x(t) \Big), \tag{2.31}$$

$$\frac{dR^z_x(t)}{dt} = \gamma I^z_x(t) + \sum_{y \in \mathcal{A}} \Big(\omega^z_{xy} R^z_y(t) - \omega^z_{yx} R^z_x(t) \Big), \tag{2.32}$$

for $z \in \mathcal{A}$. The total number of people at location x equals $N_x = \sum_z \Big(S^z_x(t) + I^z_x(t) + R^z_x(t) \Big)$. If ω^z_{xy} is independent of z, these equations reduce to the system described earlier (2.25), (2.26), and (2.27).

If the mobility rates are large compared to the epidemic parameters (infection and recovery rate), $\omega^z_{xy} \gg \beta, \gamma$, then the steady-states for the number of susceptible, infectious, and removed individuals are reached separately. This happens because in this case, the time scale of commuting $\tau_c \sim 1/\omega^z_{xy}$ is much shorter than the timescales at which people get infected or recover, $\tau_i \sim 1/\beta$ and $\tau_r \sim 1/\gamma$. This means the terms with mobility rates in (2.30), (2.31), and (2.32) vanish, and the model has reduced to the remote-force-of-infection model. Results in the remainder of this work are based on a remote-force-of-infection model.

2.2.3 Remote-force-of-infection model

The advantage of remote-force-of-infection models is that the movement of individuals is not explicitly modeled.

In the last part of the previous section, we argued that the mobility terms vanish from (2.30), (2.31), and (2.32) if we assume mobility rates to be much larger than the infection and recovery rates. In other words, the duration of visits is short compared to the timescales of the epidemic (e.g., the infectious period).

We sum (2.30), (2.31), and (2.32) over $z \in \mathcal{A}$. We obtain the following equations:

$$\frac{dS_x(t)}{dt} = -\beta \frac{S_x(t)}{N_x} \sum_{y \in \mathcal{A}} I_x^y(t), \tag{2.33}$$

$$\frac{dI_x(t)}{dt} = \beta \frac{S_x(t)}{N_x} \sum_{y \in \mathcal{A}} I_x^y(t) - \gamma I_x(t), \tag{2.34}$$

$$\frac{dR_x(t)}{dt} = \gamma I_x(t). \tag{2.35}$$

$I_x^y(t)$ are the number of infectious individuals who are currently at x but have y as their base location. $I_x^y(t)$ is a fraction of the total number of infectious individuals with base location y, and the fraction depends on the connection between location x and y. We include this fraction in the infection rate β and write β_{xy}. So, we can write in terms of the total number of infectious individuals $I_y(t)$:

$$\frac{dS_x(t)}{dt} = -\frac{S_x(t)}{N_x} \sum_{y \in \mathcal{A}} \beta_{xy} I_y(t), \tag{2.36}$$

$$\frac{dI_x(t)}{dt} = \frac{S_x(t)}{N_x} \sum_{y \in \mathcal{A}} \beta_{xy} I_y(t) - \gamma I_x(t), \tag{2.37}$$

$$\frac{dR_x(t)}{dt} = \gamma I_x(t). \tag{2.38}$$

These equations are exactly equal to remote-force-of-infection models used by others [52, 53]. This system of coupled equations has the following conservation law:

$$N_x = S_x(t) + I_x(t) + R_x(t) \quad \text{for } x \in \mathcal{A}, \tag{2.39}$$

and initial conditions $S_x(0)$, $I_x(0)$, and $R_x(0)$ for $x \in \mathcal{A}$.

In this work, we focus on a particular form of the infection rate β_{xy}. We distinguish the local transmission, β_{xx} from the non-local transmission, β_{xy} for $x \neq y$.

For local transmission, we use an infection rate β_l. For non-local transmissions, we use another infection rate β_m multiplied by a fraction that depends on how strong location x and y are coupled. In other words, it depends on how much mobility there is between x and y.

For non-local transmission, we distinguish two mechanisms: (1) infectious visitors from other locations come here and infect susceptible here, or (2) because susceptible individuals visit other regions and get infected there. Higher-order interactions are usually negligible [54].

To include these two mechanisms in β_{xy}, we introduce the mobility parameter M_{xy}. It is defined as the number of individuals who live in y and commute with location x on a daily basis. We scale M_{xy} with the population size at y, $m_{xy} = M_{xy}/N_y$. m_{xy} can be interpreted as the

probability of an individual living in y visiting location x. Note that $m_{xx} = 0$. We can estimate M_{xy} from the mobility data of The Netherlands; see Section 5.2.3.

This means that the local incidence (see Section 2.1.2) at location x equals $\beta_l S_x(t) I_x(t)/N_x$. The non-local incidence is related to mobile individuals comes in two different ways: (1) $\beta_m \frac{S_x(t)}{N_x} m_{xy} I_y(t)$ and (2) $\beta_m \frac{I_y(t)}{N_y} m_{yx} S_x(t)$. Note, m_{xy} does not need to be equal to m_{yx}.

In summary, β_{xy} as used in (2.36), (2.37), and (2.38), becomes:

$$\begin{cases} \beta_{xy} = \beta_l & \text{if } x = y, \\ \beta_{xy} = (m_{xy} + m_{yx})\beta_m & \text{if } x \neq y, \end{cases} \tag{2.40}$$

where β_l is the local transmission rate, β_m the non-local (mobile) transmission rate, and m_{xy} is the fraction of individuals with base locations y who visit location x.

When we substitute the expressions for β_{xy} in (2.36), (2.37), and (2.38) and write the equation in terms of concentrations ($s_x(t)$ instead of $S_x(t)$, etc.) we find:

$$\frac{ds_x(t)}{dt} = -\beta_l s_x(t) i_x(t) - \beta_m \sum_{y \in \mathcal{A}} (s_x(t) m_{yx} i_y(t) + i_y(t) m_{yx} s_x(t)) \tag{2.41}$$

$$\frac{di_x(t)}{dt} = \beta_l s_x(t) i_x(t) + \beta_m \sum_{y \in \mathcal{A}} (s_x(t) m_{yx} i_y(t) + i_y(t) m_{yx} s_x(t)) - \gamma i_x(t) \tag{2.42}$$

$$\frac{dr_x(t)}{dt} = \gamma i_x(t) \tag{2.43}$$

For the SEIR model, we can follow similar steps as we did so far for the SIR model. This leads to the following system of equations for the SEIR model:

$$\frac{ds_x(t)}{dt} = -\beta_l s_x(t) i_x(t) - \beta_m \sum_{y \in \mathcal{A}} \big(s_x(t) m_{yx} i_y(t) + i_y(t) m_{yx} s_x(t)\big) \tag{2.44}$$

$$\frac{de_x(t)}{dt} = \beta_l s_x(t) i_x(t) + \beta_m \sum_{y \in \mathcal{A}} \big(s_x(t) m_{yx} i_y(t) + i_y(t) m_{yx} s_x(t)\big) - \nu e_x(t) \tag{2.45}$$

$$\frac{di_x(t)}{dt} = \nu e_x(t) - \gamma i_x(t) \tag{2.46}$$

$$\frac{dr_x(t)}{dt} = \gamma i_x(t) \tag{2.47}$$

Note that in both models we sum over all other location $y \in \mathcal{A}$. When there is no edge between x and y in the meta-population network, then m_{xy} and m_{yx} are both equal to zero.

2.2.4 Epidemic threshold

As we now have many coupled subpopulations, calculating the $\mathcal{R}_0$ is not trivial. It might depend on the mobility between different locations.

To calculate the basic reproduction number, we introduce the idea of a *next generation matrix*, NGM, which we denote by $\mathbf{K}$ [45]. Elements of $\mathbf{K}$ represent the number of expected new infections caused by current infectious individuals at different locations. We can think of this as being the next generation of infectious individuals.

The NGM can be constructed using epidemiological reasoning [55]. The element in column and row x represents the average number of infections in subpopulation x caused by an infectious individual who belongs to population x. In the same column, but in row y, the element represents infections caused in population y by a contagious individual from x. In general, element K_{yx} represents infections caused in a completely susceptible population y by and an infectious individual from population x.

The number of subpopulations n determines the size of the NGM; it is a square $n \times n$ matrix. Diagonal elements of the NGM are local infections. These elements are the product of the local infection rate β_l multiplied with the infectious period $1/\gamma$. Therefore, we have,

$$K_{xx} = \frac{\beta_l}{\gamma}. \tag{2.48}$$

For off-diagonal elements, infections are related to mobility. It is the non-local infection rate β_m times the infectious period $1/\gamma$ and now multiplied with the value that someone from x going to y or the other way around, which is equal to $m_{yx} + m_{yx}$. We sum m_{xy} and m_{yx} because, as mentioned before, non-local transmission from a contagious individual from x can occur at x or at y. We find,

$$K_{xy} = (m_{xy} + m_{yx})\frac{\beta_m}{\gamma} = m'_{xy}\frac{\beta_m}{\gamma}. \tag{2.49}$$

We have seen in Section 2.1.3 that the addition of the exposed compartment does not influence the reproduction number. Similarly, here, we can follow the reasoning from the SIR model to find that elements of $\mathbf{K}$ for the SEIR model are equal to those of the SIR model.

How many new infections a specific person is expected to cause depends on how long he or she is infectious (= infectious period = $1/\gamma$) and the number of infections he or she causes per unit of time (= the infection rate β_{xy}). It depends not on whether we include an exposed compartment.

Once all elements of $\mathbf{K}$ are calculated, $\mathcal{R}_0$ is the largest eigenvalue of the NGM, denoted by $\rho(\mathbf{K})$. To understand this, consider the size of the infectious population after m generations. The infectious population is a vector of length n, $\vec{i}_0 = \big(i_1(t_0), i_2(t_0), ..., i_n(t_0)\big)$. After the m generations, $\vec{i}_m = \mathbf{K}^m \cdot \vec{i}_0$. On average, the infectious population grows with a factor $||\mathbf{K}^m||^{\frac{1}{m}}$ per generation. When m goes to infinity, we have [56]

$$\lim_{m\to\infty} ||\mathbf{K}^m||^{\frac{1}{m}} = \rho(\mathbf{K}) = \mathcal{R}_0. \tag{2.50}$$

Similarl to what we discussed in Section 2.1.4, a value of $\mathcal{R}_0 > 1$ is necessary for exponential growth in the number of infectious individuals. For a metapopulation this means that the disease can spreads through the different subpopulation before it dies out.

2.3 Stochasticity

Until now, we have only discussed models in terms of systems of differential equations. Evaluation of these systems results in a mean-field approach and is valid when the population sizes become infinitely large, often referred to as the *fluid limit* or the *thermodynamic limit*. For small

systems, stochasticity plays an important role [57, 58]. In this section, we introduce a possible way of including stochasticity into the SEIR and SIR models.

When neglecting stochastic effects, an epidemic always shows exponential growth if $\mathcal{R}_0 > 1$ initially. However, there is the probability that an infectious individual does not transmit the disease before he or she recovers while $\mathcal{R}_0 > 1$. If this happens with all infected individuals, the epidemic dies out. The probability becomes very small as the number of infectious individuals becomes large and when $\mathcal{R}_0 \gg 1$ [11].

Secondly, stochastic effects play a role in distinguishing two different kinds of infectious individuals based on their contagiousness. E.g. with a small probability, an individual is highly contagious. In a deterministic setting, we would obtain mean-field results, and it becomes impossible to see the effect of this distinction. For small numbers of infectious individuals and a distinction between their level of contagiousness, stochastic effects play an important role; see Section 5.3.

We use *stochastic rounding* as a way to include stochastic effects [46]. It is based on a forward Euler scheme that can provide a numerical solution for a differential equation with initial conditions. For example, if we want to find a numerical solution of a function $f(t)$, the forward Euler scheme is

$$f(t + \Delta t) = f(t) + \Delta t \frac{df}{dt}. \tag{2.51}$$

Here,$df(t)/dt$ is the derivative of $f(t)$ that we want the numerically evaluate. Δt is the step size. When applying stochastic rounding, we let $\Delta t \frac{df}{dt}$ be the parameter of a Poisson distribution [59]. We define the random variable $F \sim \text{Pois}(\lambda = \Delta t \frac{df}{dt})$. This results in the following scheme:

$$f(t + \Delta t) = f(t) + F. \tag{2.52}$$

In the remainder of this work, we mainly focus on a deterministic approach instead of a stochastic approach. However, in Section 5.3 and Appendix H, we do stochastic simulations.

Chapter 3

Analytical results

In this chapter, we present the analytical results for an SIR and SEIR metapopulation model. We consider a one-dimensional network, which we show in Figure 3.1. This network structure allows us to rewrite the model equations in Section 2.2.3 in terms of a continuous position coordinate instead of a discrete subpopulation number. $s_x(t) \to s(x,t)$, $i_x(t) \to i(x,t)$, $r_x(t) \to r(x,t)$, and in the SEIR model we also have $e_x(t) \to e(x,t)$

Our numerical results for localized initial conditions show different behavior for $\mathcal{R}_0 > 1$ and $\mathcal{R} \leq 1$, see Figure 3.2. For $\mathcal{R}_0 > 1$, we observe a growth in the infectious concentration $i(x,t)$, which broadens diffusively. While for $\mathcal{R}_0 \leq 1$, we observe that the infected concentration $i(x,t)$ broadens and decays exponentially. This decay of $i(x,t)$ is in line with what we saw in Figures 2.9 and Figure 2.9(b) for $i(t)$.

If $\mathcal{R}_0 > 1$, we observe that $i(x,t)$ eventually transitions into two traveling waves that move with constant shape and velocity in opposite directions in both the SIR and SEIR models. These traveling waves share properties with solitary waves [60]. In the literature, work has been done on solitary-like waves in reaction-diffusion epidemic models [61], but, to our knowledge, no reports are made for remote-force-of-infection epidemic models, which we use.

In the next chapter, we further analyze traveling waves when $\mathcal{R}_0 > 1$. The main difference compared to solitary waves is that our traveling waves annihilate during a collision with other traveling waves. We numerically evaluate this by using periodic boundaries and a finite spatial domain. Analytically, we let the spatial domain range from $-\infty$ to $+\infty$.

In Figure 3.3, we zoom in on numerical results for $\mathcal{R}_0 = 1$. We observe that $i(x,t)$ transitions into two decaying traveling waves that move in opposite directions in *both* the SIR and SEIR models. Also, we observe this for $\mathcal{R}_0 < 1$, where the waves decay even faster.

In this chapter, we derive the model equations for a continuous space coordinate in Section 3.1. Then, in Section 3.2 we argue that for localized initial conditions and for times shortly after an outbreak, $s(x,t) \approx 1$. This allows us to find solutions for the SIR and SEIR models. For the SIR model, we find a modified diffusion equation. For the SEIR model, we find a modified telegraph equation.

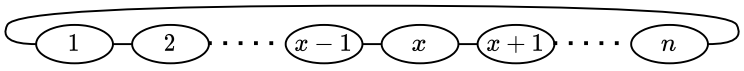


FIGURE 3.1: Metapopulation network. Each subpopulation is connected to its two nearest neighbors. Location x is an element of the set of all locations $\mathcal{A}$, $x \in \mathcal{A} = [1, ..., n]$.

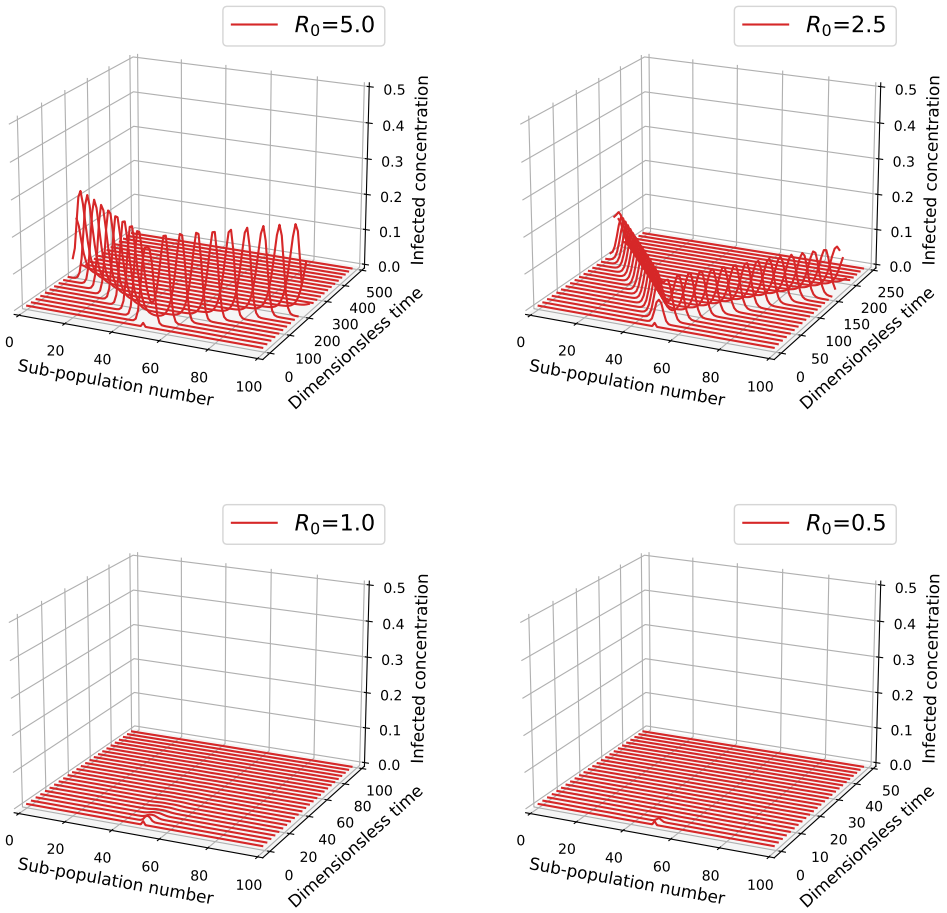


FIGURE 3.2: Numerical results of the infected concentration $i_x(t)$ for an SIR model on a network with 100 subpopulations, see Figure 3.1, for different values of $\mathcal{R}_0$. Initial conditions are $s_x(t=0)=1$, $i_x(t=0)=0$, $r_x(t=0)=0$ for $x \in [1,..,n]$, except for $x=50$, which we initialize as $s_{50}(t=0)=0.99$, $i_{50}(t=0)=0.01$, and $r_{50}(t=0)=0$.

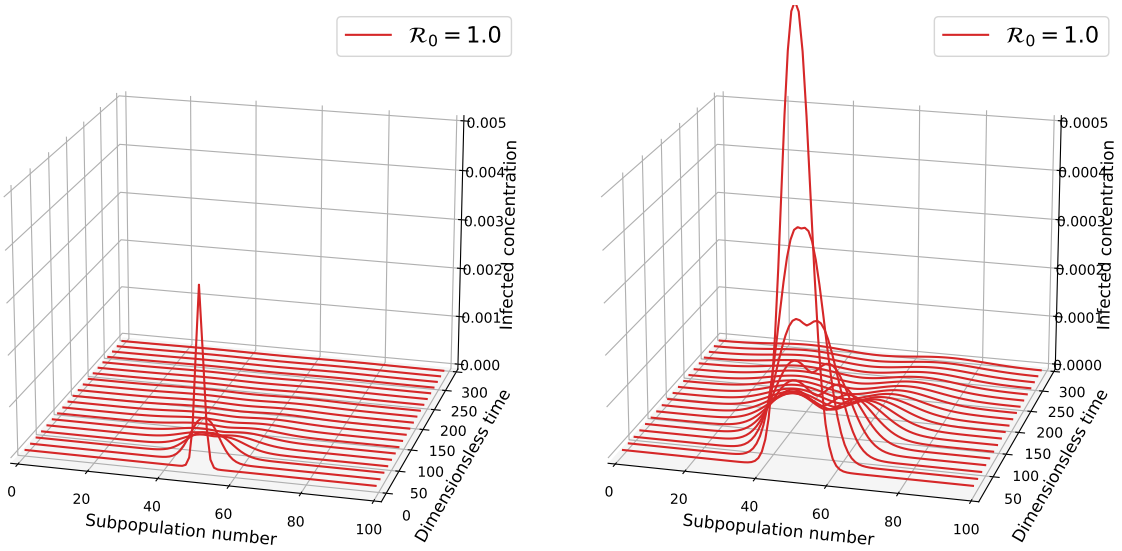


FIGURE 3.3: Numerical results of the infected concentration $i_x(t)$ for an SIR model on a network with 100 subpopulations, see Figure 3.1, for $\mathcal{R}_0 = 1$. Initial conditions are $s_x(t=0) = 1$, $i_x(t=0) = 0$, $r_x(t=0) = 0$ for $x \in [1,..,n]$, except for $x = 50$, which we initialize as $s_{50}(t=0) = 0.99$, $i_{50}(t=0) = 0.01$, and $r_{50}(t=0) = 0$. Infected concentration axes are different.

Lastly, at the wavefront, we can also assume $s(x,t) \approx 1$. We use this to find an expression for the wave propagation velocity as a function of the model parameters in Section 3.3.

The code to reproduce the figures is publicly available at https://github.com/teunhend/master-thesis.

In Chapter 4, we discuss our numerical results in more detail and compare them with our analytical results.

3.1 SIR and SEIR meta-population models

The network in Figure 3.1 allows us to rewrite the SIR model equations (2.41), (2.42), (2.43), and the SEIR model equations, (2.44), (2.45), (2.46), and (2.47), in terms of a continuous spatial coordinate.

To obtain analytical results, we make some assumptions that simplify the model equations. We assume mobility terms m_{xy} independent of x and y in both models by setting them equal to $m_{xy} \equiv m$. We assume that the population density is constant for all x. And lastly, we assume the distance between subpopulation x and $x+1$ to be equal to Δx.

In Equation (2.41) and Equation (2.42), for subpopulation x ,the sum over all connected locations y can be rewritten as a sum over populations $x-1$ and $x+1$, as we only include nearest neighbors, see Figure 3.1. Then, we approximate $i_{x-1}(t) + i_{x+1}(t)$ by using a Taylor series approximation [62]:

$$\Delta x^2 \frac{\partial^2 i(x,t)}{\partial x^2} + 2i(x,t) \approx i_{x-1}(t) + i_{x+1}(t) \tag{3.1}$$

For the SIR model, Equation (2.41) and Equation (2.42), we find the following equations:

$$\frac{\partial s(x,t)}{\partial t} = -(\beta_l + 4m\beta_m)s(x,t)i(x,t) - 2m\beta_m\Delta x^2 s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2}, \tag{3.2}$$

$$\frac{\partial i(x,t)}{\partial t} = (\beta_l + 4m\beta_m)s(x,t)i(x,t) + 2m\beta_m\Delta x^2 s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2} - \gamma i(x,t). \tag{3.3}$$

We can calculate $r(x,t)$ via the conservation law $r(x,t) = 1 - s(x,t) - i(x,t)$. We render both equations dimensionless by transforming time as $t = \frac{1}{\beta_l + 4m\beta_m}t^*$ and the spatial coordinate in our network as $x = \sqrt{\frac{2m\beta_m\Delta x^2}{\beta_l + 4m\beta_m}}x^*$. Substituting these in Equation (3.2) and Equation (3.3) and dropping the asterisks, we find:

$$\frac{\partial s(x,t)}{\partial t} = -s(x,t)i(x,t) - s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2}, \tag{3.4}$$

$$\frac{\partial i(x,t)}{\partial t} = s(x,t)i(x,t) + s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2} - \frac{1}{\mathcal{R}_0}i(x,t). \tag{3.5}$$

Here, the basic reproduction number $\mathcal{R}_0$ is the only remaining parameter and $\mathcal{R}_0 = \frac{\beta_l + 4m\beta_m}{\gamma}$. Together with initial conditions ,$s(x, t = 0)$, $i(x, t = 0)$, and $r(x, t = 0)$, and the conservation law $r(x,t) = 1 - s(x,t) - i(x,t)$, the system of PDEs in Equation (3.4) and Equation (3.5) forms a well-defined system. The domain for $t = [0, \infty)$. Note that these equations are different compared to what others use in modeling spatial spread of epidemics as they use so-called reaction-diffusion models. For a comparison, see Appendix E.

We can choose the domain of x to be infinitely large, $x = (-\infty, \infty)$, or of finite size L, $x = (-L/2, L/2)$. In the latter, we can use periodic boundary conditions. $s(x = -L/2, t) = s(x = +L/2, t)$, $i(x = -L/2, t) = i(x = +L/2, t)$, and $r(x = -L/2, t) = r(x = +L/2, t)$, or we can use absorbing boundary conditions, $s(x = \pm L/2, t) = 0$, $i(x = \pm L/2, t) = 0$, $r(x = \pm L/2, t) = 1$.

Periodic boundary conditions lead to collisions of the traveling waves, while absorbing boundary conditions lead to a collision with a wall. In both cases, we observe in numerical results that the traveling waves are annihilated.

In rewriting the SEIR model, Equations (2.44), (2.45), (2.46), and (2.47), we follow similar steps as we did for the SIR model.

We use again a Taylor series approximation to replace $i_{x-1}(t) + i_{x+1}(t) \approx \Delta x^2 \frac{d^2 i(x,t)}{dx^2} + 2i(x,t)$. Time and space in our network are transformed into dimensionless parameters via $t = \frac{1}{\beta_l + 4m\beta_m}t^*$ and $x = \sqrt{\frac{2m\beta_m\Delta x^2}{\beta_l + 4m\beta_m}}x^*$. These substitutions and transformations in combination with dropping the asterisk lead to the following equations:

$$\frac{\partial s(x,t)}{\partial t} = -s(x,t)i(x,t) - s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2} \tag{3.6}$$

$$\frac{\partial e(x,t)}{\partial t} = s(x,t)i(x,t) + s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2} - ce(x,t) \tag{3.7}$$

$$\frac{\partial i(x,t)}{\partial t} = ce(x,t) - \frac{1}{\mathcal{R}_0}i(x,t) \tag{3.8}$$

Based on Section 2.1.4, we expect $\mathcal{R}_0$ to be equal in the SIR and SEIR models. Here, the basic reproduction number $\mathcal{R}_0$ is equal to what we obtained in the SIR model. Namely, $\mathcal{R}_0 = \frac{\beta_l + 4m\beta_m}{\gamma}$. The parameter c is related to the rate at which individuals leave the exposed compartment,

$c = \frac{\nu}{\beta_l + 4m\beta_m}$. Like the latent period in Section 2.1.1, $1/\nu$, we interpret $1/c$ as the dimensionless latent period.

We calculate $r(x,t)$ using Equations (3.6), (3.7), (3.8) and the conservation law as via $r(x,t) = 1 - s(x,t) - e(x,t) - i(x,t)$. Together with initial conditions, $s(x,t=0)$, $e(x,t=0)$, $i(x,t=0)$, and $r(x,t=0)$, and the domains of $t = [0,\infty)$ and $x = [-\infty,\infty]$, Equations (3.6), (3.7), (3.8) form a well-defined system.

In discrete systems, we initialize a localized outbreak by setting the infected concentration in one subpopulation x_0 initially at i_0 and zero is all other subpopulations in the case of an SIR model:

$$s_x(t=0) = 1 - i_0\delta_{x,x_0}, \tag{3.9}$$
$$i_x(t=0) = i_0\delta_{x,x_0}. \tag{3.10}$$
$$r_x(t=0) = 0. \tag{3.11}$$

Here, δ_{x,x_0} is a Kronecker delta function that is equal to 1 when $x = x_0$ and equal to 0 otherwise. For the SEIR model, we expose a fraction in a single subpopulation to initialize a localized outbreak:

$$s_x(t=0) = 1 - e_0\delta_{x,x_0}, \tag{3.12}$$
$$e_x(t=0) = e_0\delta_{x,x_0}. \tag{3.13}$$
$$i_x(t=0) = 0. \tag{3.14}$$
$$r_x(t=0) = 0. \tag{3.15}$$

In this case, e_0 is the initial fraction of exposed individuals in subpopulation x_0.

In the continuum limit, we use a delta distribution function [24]. In the SIR model, we scale the delta peak with i_0:

$$s(x,t=0) = 1 - i_0\delta(x - x_0), \tag{3.16}$$
$$i(x,t=0) = i_0\delta(x - x_0), \tag{3.17}$$
$$r(x,t=0) = 0. \tag{3.18}$$

Here, $\delta(x - x_0)$ is a delta function, and i_0 is the initial concentration, and x_0 the spatial coordinate of the location of the delta peak. In the SEIR model, we scale the delta peak with e_0:

$$s(x,t=0) = 1 - e_0\delta(x - x_0), \tag{3.19}$$
$$e(x,t=0) = e_0\delta(x - x_0), \tag{3.20}$$
$$i(x,t=0) = 0, \tag{3.21}$$
$$r(x,t=0) = 0. \tag{3.22}$$

This approach is somewhat artificial as at $t = 0$; the delta function is of infinite height. But it ensures that the total concentration of infected individuals at $t = 0$ equals i_0 in the SIR model (or the total concentration of exposed individuals in the SEIR model):

$$\int_{-\infty}^{\infty} i_0\delta(x - x_0)dx = i_0 \tag{3.23}$$

In the next section, we derive a solution for $i(x,t)$ shortly after a localized outbreak.

3.2 Model behavior shortly after outbreak

This section presents analytical solutions for infectious concentration $i(x,t)$ for the SIR and SEIR, which describe their behavior shortly after initialization. The infectious concentration is influenced by $\mathcal{R}_0$.

For $\mathcal{R}_0 > 1$, $i(x,t)$ grows and broadens diffusively. Then, $i(x,t)$ transitions into two waves that travel in the opposite direction, see Figure 3.2. For $\mathcal{R}_0 \leq 1$, $i(x,t)$ decays and broadens diffusively. Due to the decay, at some point, the broadening stops. Figure 3.3 zooms in on the result for $\mathcal{R}_0 = 1$ from Figure 3.2. While decaying, $i(x,t)$ transitions into two decaying traveling waves that move in different directions.

In this section, we provide a solution for the diffusive regime. In the next section, we present results for the traveling wave regime (for $\mathcal{R}_0 > 1$). Here, we focus on the diffusive-like behavior.

We argue that the approximation $s(x,t) \approx 1$ is reasonable shortly after a localized outbreak. This approximation allows us to find analytical solutions asymptotically. However, in Section 2.1.4, we argued that $\mathcal{R}_0 > 1$ is necessary for exponential growth, but at time t, there is only exponential growth if the product $s(x,t)\mathcal{R}_0 > 1$. By approximating $s(x,t) = 1$, we can only find solutions that either exponentially grow or decay. Still, numerical results show that growth and decay succeed each other for $\mathcal{R}_0 > 1$ in the SIR model (and also in the SEIR model).

We start with the SIR model. Firstly, we rewrite Equation (3.4) to

$$\frac{1}{s}\frac{\partial s}{\partial t} = -\left(i + \frac{\partial^2 i}{\partial x^2}\right). \tag{3.24}$$

Integrating both sides with respect to t gives

$$\ln \frac{s(x,t)}{s(x,0)} = -\int_{t=0}^{t} (i + \frac{\partial^2 i}{\partial x^2})\partial t, \tag{3.25}$$

$$s(x,t) = s(x,0) \exp\left[-\int_{t=0}^{t} (i + \frac{\partial^2 i}{\partial x^2})dt\right]. \tag{3.26}$$

For small values of t, the exponent is close to zero and therefore $s(x,t) \approx s(x,0)$, which is equal to Equation (3.16). The value i_0 is the concentration we initialize as infectious. As we choose this value to be small, we approximate $s(x,0) \approx 1$. So, also $s(x,t) \approx 1$. We substitute this in (3.5) and find for small values of t:

$$\frac{\partial i}{\partial t} = \frac{\partial^2 i}{\partial x^2} + \left(1 - \frac{1}{\mathcal{R}_0}\right) i. \tag{3.27}$$

We solve Equation (3.27) using Green's functions. The derivation can be found in Appendix C. We use the initialization described by Equations (3.16), (3.17), and (3.18). The solution we find is very similar to the standard diffusion equation:

$$i(x,t) = \frac{i_0}{\sqrt{4\pi t}} \exp\left[-\frac{(x-x_0)^2}{4t}\right] \exp\left[\left(1 - \frac{1}{\mathcal{R}_0}\right)t\right]. \tag{3.28}$$

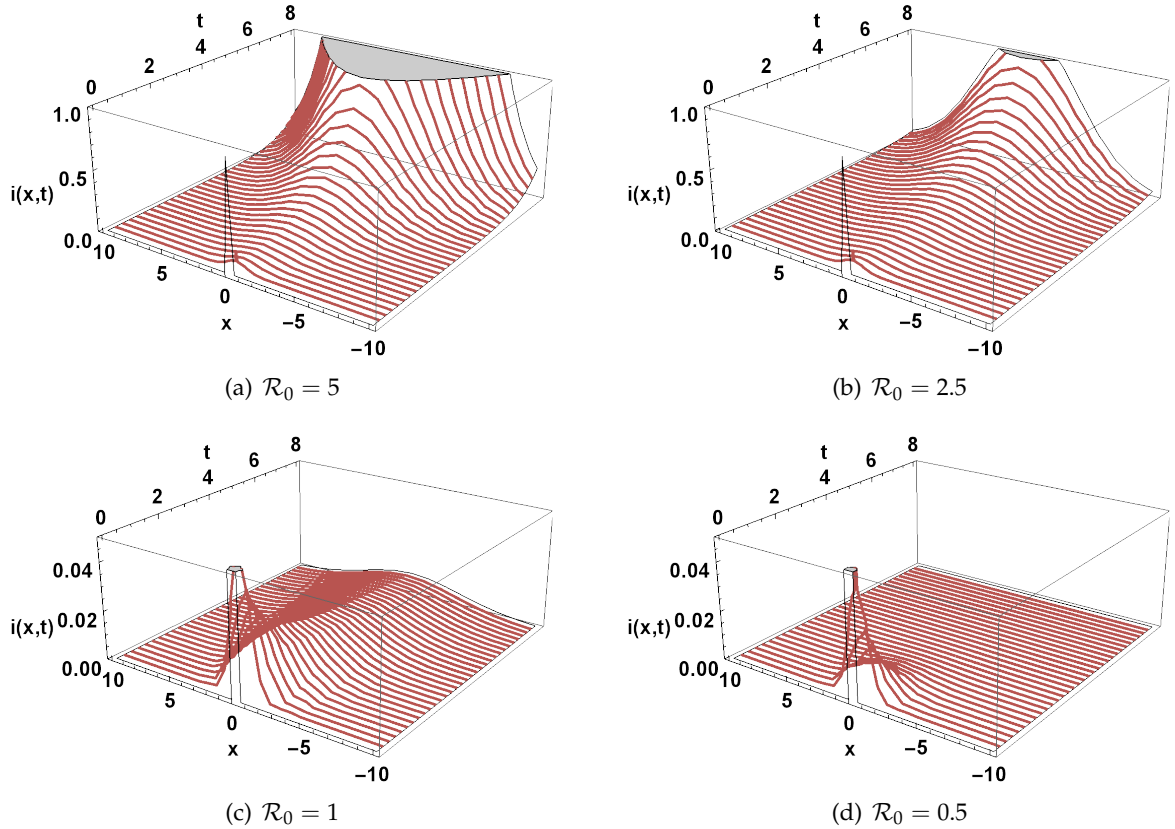
For all values of $\mathcal{R}_0$, the disease spreads diffusively. As this solution describes the PDE in (3.5) shortly after an outbreak, we conclude that what we see in numerical results is indeed diffusive behavior.

We have an extra term that depends on the basic reproduction number $\mathcal{R}_0$. Depending on the value of $\mathcal{R}_0$, the total infected concentrations grows, remains constant, or decays. For $\mathcal{R}_0 < 1$ the total infected concentration exponentially decays, while for $\mathcal{R}_0 > 1$ the solution gets exponentially larger. For $\mathcal{R}_0 = 1$, we obtain the standard diffusion equation.

Figure 3.4 shows solutions of Equation (3.28). For $\mathcal{R}_0 > 1$, we see that indeed $i(x,t)$ grows and broadens. However, $i(x,t)$ continues to grow to the point where $i(x,t) > 1$, which is not allowed as $i(x,t)$ is a concentration value and cannot be larger than unity.

For $\mathcal{R}_0 \leq 1$, we see a resemblance with Figure 3.2. However, when we take a closer look at Figure 3.3, we observe a transition into two decaying traveling waves, which we do not observe in Figure 3.4(c) and Figure 3.4(d).

Equation (3.28) describes observed behavior from Figure 3.2 shortly after the outbreak. Still, for $\mathcal{R}_0 > 1$ it cannot describe the transition into two traveling waves, nor can it describe the two decaying waves for $\mathcal{R}_0 \leq 1$. A solution that is *not* based $s(x,t) = 1$ can probably provide a solution for this transition.



(a) $\mathcal{R}_0 = 5$ (b) $\mathcal{R}_0 = 2.5$

(c) $\mathcal{R}_0 = 1$ (d) $\mathcal{R}_0 = 0.5$

FIGURE 3.4: Numerical evaluation of Equation 3.28 for different values of $\mathcal{R}_0$ and for $x_0 = 0$ and $i_0 = 0.01$.

In the case of the SEIR model, we also approximate $s(x,t) \approx 1$. This is based on similar arguments as for the SIR model. We take Equation (3.8) and calculate its derivative with respect to time:

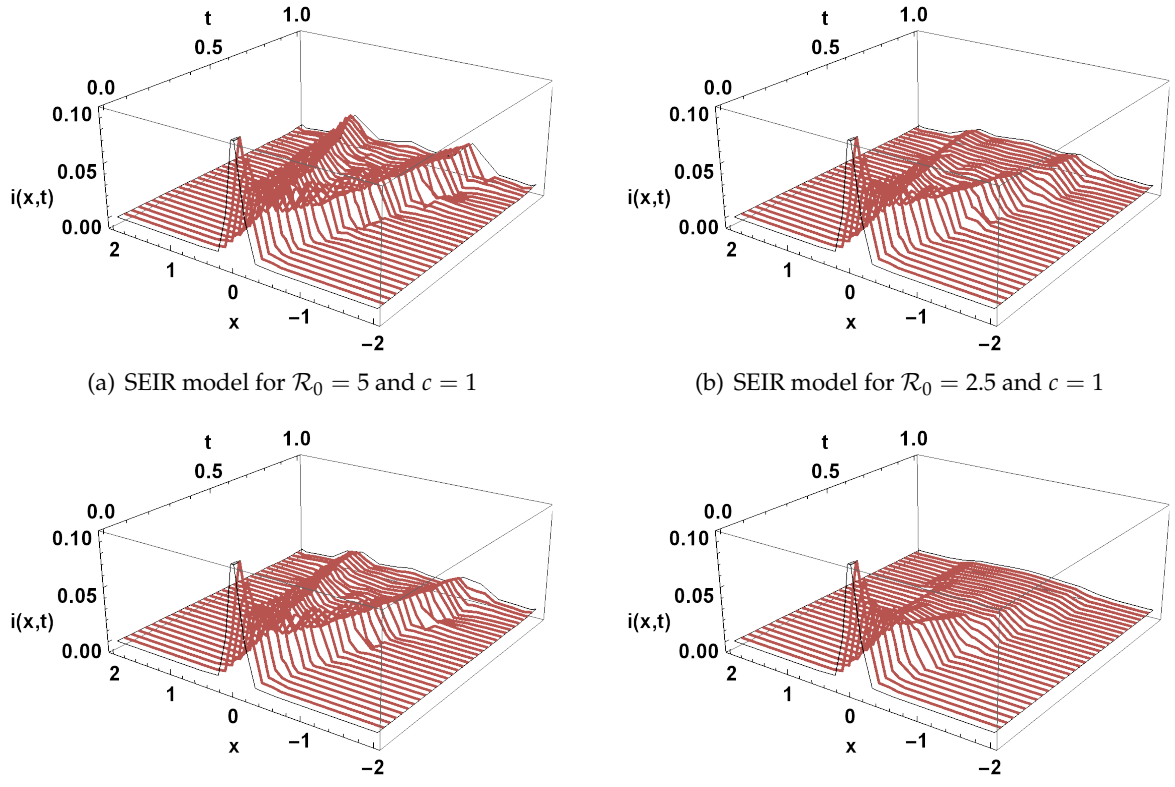
$$\frac{\partial e(x,t)}{\partial t} = \frac{1}{c}\frac{\partial^2 i(x,t)}{\partial t^2} + \frac{1}{c\mathcal{R}_0}\frac{\partial i(x,t)}{\partial t}. \tag{3.29}$$

We substitute this equation and Equation (3.8) into Equation (3.7). Together with $s(x,t) \approx 1$, we obtain an equation in terms of $i(x,t)$ and model parameters. :

$$\frac{\partial^2 i(x,t)}{\partial t^2} + \left(\frac{1}{\mathcal{R}_0} + c\right)\frac{\partial i(x,t)}{\partial t} - c\frac{\partial^2 i(x,t)}{\partial x^2} = c\left(1 - \frac{1}{\mathcal{R}_0}\right) i(x,t) \tag{3.30}$$

This equation is better known as the general *telegraph equation*. The telegraph equation has been studied extensively, and its properties are well known [63]. For small values of t, the solution shows traveling wave solutions, while for large values of t, the solution becomes diffusive.

We numerically evaluate Equation (3.30), and we show the results in Figure 3.5. While Figure 3.2 and Figure 3.3 are based on an SIR model, the SEIR model results are similar as in that they also show decaying traveling waves for $\mathcal{R}_0 \leq 1$ and diffusive growth followed by traveling waves of constant shape for $\mathcal{R}_0 > 1$. From Figure 3.5, we find that Equation (3.30) describes the SEIR model well for $\mathcal{R}_0 \leq 1$. It shows diffusive decay, then it transitions into decaying traveling waves. For $\mathcal{R}_0 > 1$, Equation (3.30) results in exponential growth of $i(x,t)$ beyond $i(x,t) = 1$ for $t \gg 1$.



(a) SEIR model for $\mathcal{R}_0 = 5$ and $c = 1$

(b) SEIR model for $\mathcal{R}_0 = 2.5$ and $c = 1$

(c) SEIR model for $\mathcal{R}_0 = 1$ and $c = 1$

(d) SEIR model for $\mathcal{R}_0 = 0.2$ and $c = 1$

FIGURE 3.5: Numerical evaluation of Equation 3.30 for different values of $\mathcal{R}_0$ and c. Initial condition is a delta peak at $x = 0$, $i(x, t = 0) = i_0\delta(x)$, where $i_0 = 0.01$.

We provide a solutions for Equation (3.30) based on a following boundary and initial conditions for the domain $x \in [0, L]$ and $t \in [0, \infty)$:

$$i(x,0) = i_0\delta(x - x_0), \qquad \left.\frac{\partial i}{\partial t}\right|_{t=0} = 0, \qquad i(0,t) = 0, \qquad i(L,t) = 0.$$

Using a method called *separation of variables* [64], we find the following solution:

$$i(x,t) = \frac{1}{2}\sum_{n=1}^{\infty} A_n \exp\left[-\left(\frac{c}{2} + \frac{1}{2\mathcal{R}_0}\right)t\right]\left(\sin\left(\frac{n\pi x}{L} - \omega_n t + \phi_n\right) + \sin\left(\frac{n\pi x}{L} + \omega_n t - \phi_n\right)\right), \tag{3.31}$$

where the constants are as follows:

$$\omega_n = \sqrt{c\left(\left(\frac{n\pi}{L}\right)^2 - \frac{5}{4} - \frac{1}{4c\mathcal{R}_0}\right)}, \tag{3.32}$$

$$\phi_n = \arctan\left(\frac{c}{2\omega_n} + \frac{1}{2\omega_n\mathcal{R}_0}\right), \tag{3.33}$$

$$A_n = \frac{2}{L\cos\phi_n}\sin\left(\frac{n\pi x_0}{L}\right). \tag{3.34}$$

In Appendix D, we present the full derivation of Equation (3.31).

From Equation (3.31), we can see that the wave-like solutions decay exponentially with an exponent dependent on c and $\mathcal{R}_0$.

3.3 Wave propagation velocity

After some time, the solutions to our dynamical equations transition into a regime where two traveling waves are observed. For $\mathcal{R}_0 > 1$, we observe traveling waves with constant velocity and shape. In Section 4.1.1, we show that the numerical results strongly point towards a temporally invariant solution, i.e., can be written as $i(x,t) = f(x + vt) + f(x - vt) \equiv i(u)$ with $u = x \pm vt$.

Here, we present analytical results for the wave velocity. The most important approximation we use throughout this section is that at the wavefront $s(x,t) \approx 1$.

SIR Model. Here, we write the solution of the SIR equations as $i(x,t) = f(x + v^*t) + f(x - v^*t) \equiv i(u)$, where v^* is the dimensionless velocity (and x and t are the dimensionless position and time) and $u = x \pm v^*t$. This allows us to rewrite the PDE of (3.5) into an ODE in terms of u via $\frac{\partial}{\partial t} \to \pm v^*\frac{\partial}{\partial u}$ and $\frac{\partial^2}{\partial x^2} \to \frac{\partial^2}{\partial u^2}$. We define $i = i(u)$, $i' = \frac{\partial i(u)}{\partial u}$ and $i'' = \frac{\partial^2 i(u)}{\partial u^2}$ and obtain:

$$0 = si'' + v^*i' + \left(s - \frac{1}{\mathcal{R}_0}\right)i. \tag{3.35}$$

For sufficiently localized initialization ($s(x,t) \approx 1$ at the wavefront), the PDE reduces to an ODE:

$$0 = i'' + v^*i' + \left(1 - \frac{1}{\mathcal{R}_0}\right)i. \tag{3.36}$$

We know the solution of this equation is of the following form:

$$i(u) = c_1 \exp\left[\lambda_1 u\right] + c_2 \exp\left[\lambda_2 u\right]. \tag{3.37}$$

Here, c_1 and c_2 are constants determined by initial conditions, and λ_1 and λ_2 are solutions to the so-called *characteristic equation*:

$$0 = \lambda^2 + v^*\lambda + \left(1 - \frac{1}{\mathcal{R}_0}\right). \tag{3.38}$$

$i(u)$ has traveling wave solutions for real roots of Equation (3.38). This means that only for specific values of v^*, traveling waves can be observed:

$$v^* \geq 2\sqrt{\frac{\mathcal{R}_0 - 1}{\mathcal{R}_0}}. \tag{3.39}$$

For $\mathcal{R}_0 \to \infty$, the dimensionless velocity obtains its maximum $v^* \to 2$. This equation only holds when $\mathcal{R}_0 > 1$, as only then we observe traveling waves with constant shape and velocity. We can transform this dimensionless velocity to a velocity with dimensions, using the position and time scaling constants. This results in:

$$v \geq 2\sqrt{2m\beta_m \Delta x^2(\beta_l + 4m\beta_m - \gamma)}. \tag{3.40}$$

This velocity is based on our SIR model, which is from the class of remote-force-of-infection models. Surprisingly, this result is equal to the results obtained for reaction-diffusion epidemic models [65, 66]. We refer to Appendix E for the differences between these two classes of models.

The actually observed velocity turns out to be equal to the lowest possible value of v^* for reaction-diffusion models [65]. For localized initial conditions, our wavefront is a so-called *pulled* wavefront, for which the relevant propagation velocity is the minimal velocity [67]. We see in Section 4.1.2 that for our remote-force-of-infection models, the observed velocity is also the lowest possible value of v^*.

Following similar steps as in our discussion of the SIR model, we find the following ODE for the SEIR model:

$$0 = \left(1 - \frac{(v^*)^2}{c}\right) i'' + v^* \left(c + \frac{\mathcal{R}_0}{c}\right) i' + (1 - \mathcal{R}_0)\, i. \tag{3.41}$$

Again, we use the characteristic equation to find only real roots, which puts a condition on v^*. For sufficiently localized initialization, this system has traveling wave solutions for

$$v^* \geq 2\sqrt{\frac{\mathcal{R}_0 - 1}{\mathcal{R}_0} \frac{c^2}{c^2 + 4(c - 2/\mathcal{R}_0) + (1/\mathcal{R}_0)^2}}. \tag{3.42}$$

For $c \to \infty$, (3.42) reduces to (3.39), just as expected. Because $1/c$ can be interpreted as the dimensionless latent period, the time spent in the exposed compartment. For $c \to \infty$, the time spent in the exposed compartment goes to zero, reducing the SEIR model to the SIR model.

We compare the dimensionless velocities v^* from Equations (3.39) and (3.42) in Figure 3.6. We observe that for large values of $\mathcal{R}_0$ the system obtains its maximum velocity. For the SIR mode, this maximum velocity equals 2. For the SEIR model, the maximum is determined by the model parameter c. For small dimensionless latent periods, $1/c$, or in other words, if the time spent in the exposed compartment is short, the velocity is high. For large values of $\mathcal{R}_0$, Equation (3.42) reduces to:

$$v^* = \pm 2\sqrt{\frac{c}{c+4}}. \tag{3.43}$$

In Section 4.1.2, we compare numerically obtained velocities in the SIR and SEIR models to the velocities from Equations (3.39) and (3.42). Similar to what has been observed for epidemic reaction-diffusion models, the wave propagation velocity equals to the lowest possible value of Equations (3.39) and (3.39). We show these values in Figure 3.6.

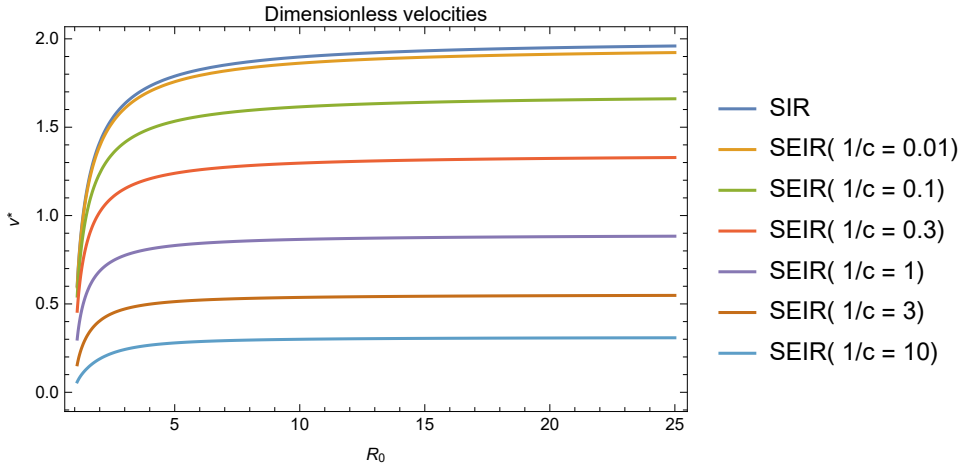


FIGURE 3.6: Analytical wave propagation velocities v^* for the SIR model and SEIR model as a function of the basic reproduction number $\mathcal{R}_0$. For the SIR model, we show Equation (3.39). For the SEIR model, we show Equation (3.42) for different values of $1/c$. $1/c$ is the dimensionless latent period, the time spent in the exposed compartment.

3.4 Other meta-population models

In Section 2.1.1, we described different epidemic compartmental models. In this chapter so far, we analyzed the SIR and the SEIR model. For the reader more interested in the other compartmental models, we summarize the analytical results in this section. The models that we consider are the SIS (*susceptible-infected-susceptible*) model, the SIRS (*susceptible-infected-recovered-susceptible*) model, and the SEIRS (*susceptible-exposed-infected-recovered-susceptible*) model. We leave the numerical analysis of the SIS, SIRS, SEIRS models to future work. In the next chapter, we numerically analyze the SIR and the SEIR models.

Following similar steps as in the previous sections of this chapter, we derive model equations using a continuous spatial coordinate for a network structure as depicted, in Figure 3.1. We show the model equation and relate these to asymptotic solutions, which are valid shortly after an outbreak. Also, we briefly present wave propagation velocities.

For results shortly after a localized outbreak, we see that it does not matter whether or not resistance is eventually lost. What matters is whether or not individuals become directly infectious instead of becoming exposed.

SIS model. People become susceptible to the disease after they have been infectious. In Equations (A.5) and (A.6), non-local transmission is included, similar to (2.41). The result is then transformed using a dimensionless time and dimensionless position, which results in:

$$\frac{\partial s(x,t)}{\partial t} = -s(x,t)i(x,t) - s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2} + \frac{1}{\mathcal{R}_0}i(x,t), \tag{3.44}$$

$$\frac{\partial i(x,t)}{\partial t} = s(x,t)i(x,t) + s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2} - bi(x,t) = -\frac{\partial s(x,t)}{\partial t}. \tag{3.45}$$

Here, $\mathcal{R}_0$ is the only remaining parameter, and $\mathcal{R}_0 = \frac{\beta_l + 4m\beta_m}{\gamma}$. Together with initial conditions $s(x,0)$ and $i(x,0)$ and the conservation equation $1 = s + i$, the system of PDEs in Equations (3.44) and (3.45) form a well-defined system. Using $s = 1 - i$, we can reduce the system to a single PDE,

$$\frac{\partial i(x,t)}{\partial t} = (1 - i(x,t) - b)i(x,t) + (1 - i(x,t))\frac{\partial^2 i(x,t)}{\partial x^2}. \tag{3.46}$$

For $i(x,t)$ small, Equation (3.46) reduces to Equation (3.27). This means that the initial response to a localized outbreak in the SIR and SIS model are described by the same asymptotic solution in Equation (3.28). This also means for the SIS model that the wave propagation velocity is expected to be equal to Equation 3.39, if $\mathcal{R}_0 > 1$.

SIRS model. Following similar steps as described above, we find the dimensionless equation for the meta-population SIRS model to be

$$\frac{\partial s(x,t)}{\partial t} = -s(x,t)i(x,t) - s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2} + d(1 - s(x,t) - i(x,t)), \tag{3.47}$$

$$\frac{\partial i(x,t)}{\partial t} = s(x,t)i(x,t) + s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2} - \frac{1}{\mathcal{R}_0}i(x,t). \tag{3.48}$$

Here, $\mathcal{R}_0 = \frac{\beta_l + 4m\beta_m}{\gamma}$ and $d = \frac{\alpha}{\beta_l + 4m\beta_m}$. α equals the rate at which individuals lose their immunity to the disease. Together with initial conditions $s(x,0)$ and $i(x,0)$ and the conservation equation $r = 1 - s - i$, the system of PDEs is a well-defined system. Note that for $d \to \infty$ this model reduces to the SIS model, while for $d \to 0$, this model reduces to the SIR model.

For $i(x,t)$ small, Equation (3.48) reduces to Equation (3.27). This means that the initial response to a localized outbreak in the SIR and SIRS model is described by the same asymptotic solution in Equation (3.28). Also, for the SIRS model, the wave propagation velocity is expected to be equal to Equation 3.39, if $\mathcal{R}_0 > 1$.

SEIRS model. Following similar steps as described above, we find for the dimensionless equation for the meta-population SIRS model:

$$\frac{\partial s(x,t)}{\partial t} = -s(x,t)i(x,t) - s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2} + d(1 - s(x,t) - i(x,t) - e(x,t)), \tag{3.49}$$

$$\frac{\partial e(x,t)}{\partial t} = s(x,t)i(x,t) + s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2} - ce(x,t), \tag{3.50}$$

$$\frac{\partial i(x,t)}{\partial t} = ce(x,t) - \frac{1}{\mathcal{R}_0}i(x,t), \tag{3.51}$$

where $\mathcal{R}_0 = \frac{\beta_l + 4m\beta_m}{\gamma}$, $c = \frac{\nu}{\beta_l + 4m\beta_m}$, and $d = \frac{\alpha}{\beta_l + 4m\beta_m}$. Again, α equals the rate at which individuals lose their immunity to the disease. Together with initial conditions $s(x,0)$, $e(x,0)$, and $i(x,0)$ and the conservation equation $r = 1 - s - e - i$, the system of PDEs is a well-defined system. Note, for $d \to \infty$, this model reduces to the SEIS model (not discussed here), while for $d \to 0$, this model reduces to the SEIR model.

For $i(x,t)$ small, Equation (3.51) reduces to Equation (3.30). This means that the initial response to a localized outbreak in the SEIR and SEIRS model are described by the same asymptotic solution in Equation (3.31). Also, for the SEIRS model, the wave propagation velocity is expected to be equal to Equation (3.42), if $\mathcal{R}_0 > 1$.

We leave the numerical analysis of the SIS, SIRS, SEIRS models to future work. In the next chapter, we numerically analyze the SIR and the SEIR models.

Chapter 4

Numerical results

In this chapter, we present and analyze numerical results for the SIR and SEIR models. In the previous chapter, we showed that for a specific metapopulation structure, as in Figure 3.1, we can rewrite $s_x(t) \rightarrow s(x,t)$, $e_x(t) \rightarrow e(x,t)$, etc.

For this one-dimensional network, we discuss the traveling waves that appear for $\mathcal{R}_0 > 1$ in both the SIR and SEIR models. We show arguments why the solution in the traveling wave regime is an invariant solution, i.e. $i(x,t) = i(x \pm v^* t) \equiv i(u)$. In addition, we analyze how the model parameters influence the shape of the $i(x,t)$ curve, we compare the SIR model to the SEIR model, and we verify the dimensionless velocity equations (3.39) and (3.42). We find that the traveling waves are robust to parameter variations, population density variations, and stochastic effects.

Numerically, we also analyze other metapopulation networks. Such as a network that resembles a two-dimensional grid. Subpopulations are connected to their four or eight nearest neighbors. We observe similar behavior as in our one-dimensional network. Namely, a diffusive decay for $\mathcal{R}_0 \leq 1$ and diffusive growth for $\mathcal{R}_0 > 1$, which transitions into a circular traveling wavefront. Even for multiple initial outbreak locations, a large circular wavefront appears eventually.

For more complex metapopulation structures where subpopulations are not connected to only their nearest neighbors, wave-like patterns are also observed [8]. We expect that our results from Chapter 3 and Chapter 4 also hold for more complex metapopulation structures. In Chapter 5, such a complex metapopulation structure is considered, as it is based on mobility patterns of The Netherlands.

4.1 Epidemic waves

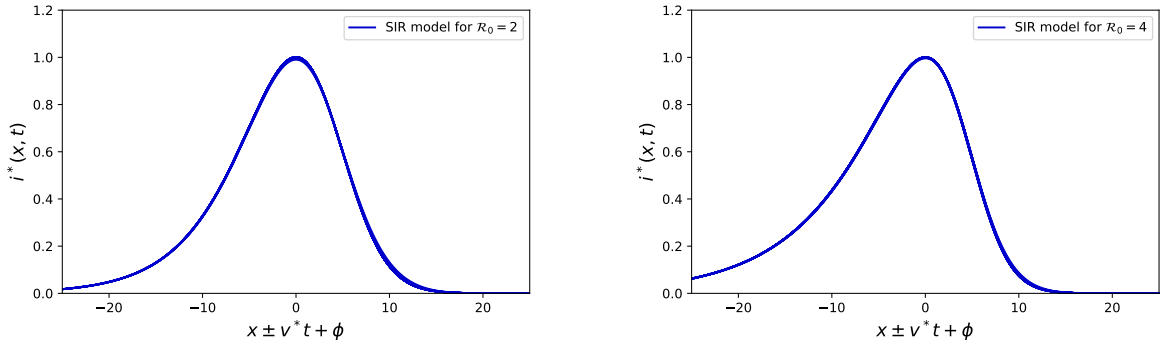
Firstly, we show that within the traveling-wave regime, $\mathcal{R}_0 > 1$, the wave's shape remains constant for a given combination of model parameters. Therefore, the solution can be written as $i(x,t) = i(x \pm vt)$. To demonstrate this, we use a network similar to Figure 3.1 with $n = 100$. Then, we discuss the influence of the model parameter on the actual shape of $i(x,t)$. In the last part of this section, we present numerical results for the propagation velocity of the wave.

In Figure 4.1, we present the infected concentration $i(x,t)$ in the SIR model for two values of $\mathcal{R}_0$. $i(x,t)$ is scaled by the peak concentration $i_{\max}$. We define the normalized infected concentration $i^*(x,t)$ as $i(x,t)/i_{\max}$. We plot $i^*(x,t)$ versus $x \pm v^* t$, where x is the dimensionless positions, v^* the dimensionless velocity (see Equation (3.39)), and t the dimensionless time. We add ϕ, such that the peak of $i(x,t)$ is at exactly zero.

The spatial domain $x \in [0, 1, ..., 99]$ and we initialize as described by Equations (3.9), (3.10), (3.11). We arbitrarily set $x_0 = 50$ and $i_0 = 0.01$. Results in this section do not depend on the

exact value of i_0. The system has periodic boundaries. So, eventually, the waves collide at $x = 0$. All positions $x \in [1, 2, .., 49]$ (wave traveling to the left) and $x \in [51, 52, ..., 99]$ (wave traveling to the right) are included in Figure 4.1. The collapse of all $i^*(x,t)$ on a single curve strongly hints towards a traveling wave solution with a constant propagation velocity.

Similarly, for the SEIR model, we present the normalized infected concentration $i^*(x,t)$ versus $x \pm v^*t$ in Figure 4.2. Here, v^* is based on Equation (3.42). We plot $i^*(x,t)$ for two different values of $\mathcal{R}_0$ and c. Again, the collapse of all $i^*(x,t)$ on a single curve strongly hints towards a traveling wave solution with a constant propagation velocity.



(a) (b)

FIGURE 4.1: Scaled infected concentration $i^*(x,t)$ in the SIR model as a function of $x \pm v^*t + \phi$ for two values of $\mathcal{R}_0$. The infected concentration $i(x,t)$ is scaled such that maximum is equal to 1. v^* is the dimensionless velocity as described by Equation (3.39). ϕ is added such that the peak of the curve is at 0. Left and right show the collapse of $i^*(x,t)$ onto a single curve.

Solitary waves are waves that maintain their shape while they propagate with constant velocity [60]. While our traveling waves appear similar to solitary waves, our traveling waves are slightly different. Like solitary waves, our traveling waves are localized and of constant shape and velocity, but they behave differently in a collision compared to solitary waves.

We observe that our traveling waves are annihilated after a collision. We show this in Figure 3.2, where for $\mathcal{R}_0 = 5$, the traveling waves collide at around $t = 400$ and annihilate. The waves collide because the system has periodic boundaries. This is completely different than what we would expect of a solitary wave. A solitary wave survives a collision with other solitary waves without losing energy and emerges from a collision with only a phase shift [60].

We analyze the shape of the normalized curves in Figure 4.1 and Figure 4.2 in the next section. In particular, we show results for a universal curve, results on the max curve height $i_{\max}$, and results for the width of the curves via the full width at half maximum, FWHM.

4.1.1 Shape and parameter influence

Based on Figure 4.1 and Figure 4.2, we try to find a universal curve. For the SIR model, we see that higher values of $\mathcal{R}_0$ result in a more asymmetric shape of $i^*(x,t)$. This relation is confirmed in Figure 4.3. For the SEIR model, we see that the shape is influenced by $\mathcal{R}_0$ in in a completely different way in Figure 4.4. Larger values of $\mathcal{R}_0$, reduce the width of the curve, instead of increasing the asymmetry. When we compare multiple ratios of the latent period over the infectious period, $1/(c\mathcal{R}_0)$, we see that the shapes of $i^*(x,t)$ overlap.

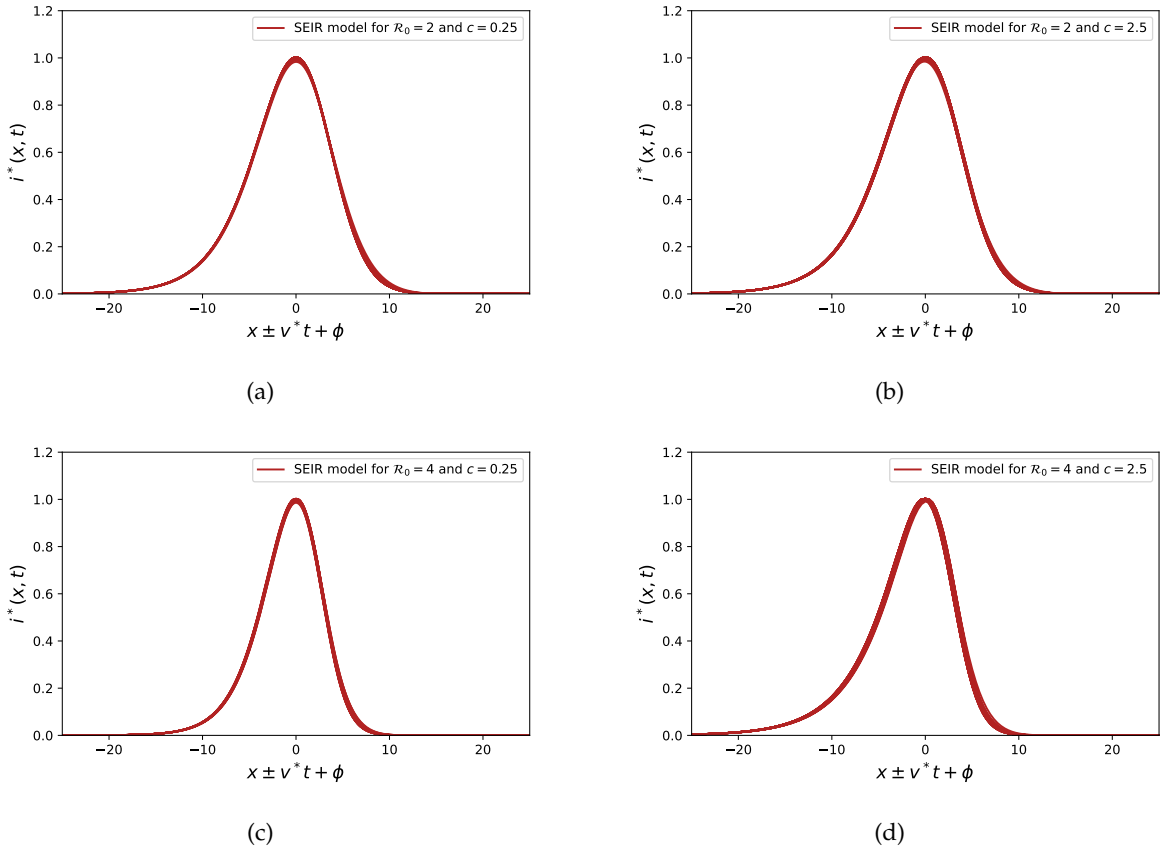


FIGURE 4.2: Scaled infected concentration $i^*(x,t)$ in the SEIR model as a function of $x \pm v^* t + \phi$ for different values of $\mathcal{R}_0$ and c. The infected concentration $i(x,t)$ is scaled such that maximum is equal to 1. v^* is the dimensionless velocity as described by Equation (3.42). ϕ is added such that the peak of the curve is at 0. Left and right show the collapse of $i^*(x,t)$ onto a single curve.

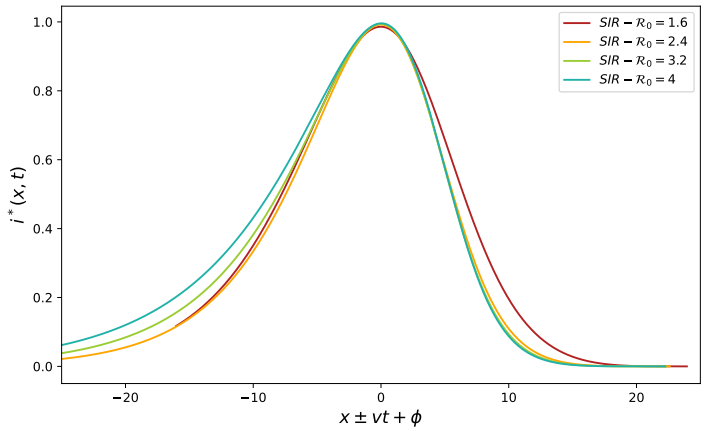


FIGURE 4.3: Scaled infected concentration $i^*(x,t)$ in the SIR model as a function of $x \pm v^* t + \phi$ for multiple values of $\mathcal{R}_0$ in a single figure. v^* is the dimensionless velocity as described by Equation (3.39). ϕ is added such that the peak of the curve is at 0.

Before we show results on the maximum height of the non-normalized curve $i(x,t)$, $i_{\max}$, we briefly recapitulate on the *final size of the epidemic* r_∞. In Section 2.1.4, we discussed the final

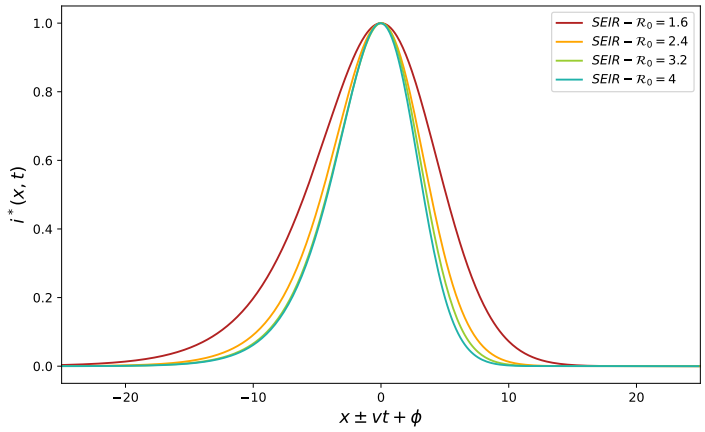


FIGURE 4.4: Scaled infected concentration $i^*(x,t)$ in the SEIR model as a function of $x \pm v^*t + \phi$ for multiple values of $\mathcal{R}_0$ in a single figure while $1/c = 1$. v^* is the dimensionless velocity as described by Equation (3.42). ϕ is added such that the peak of the curve is at 0.

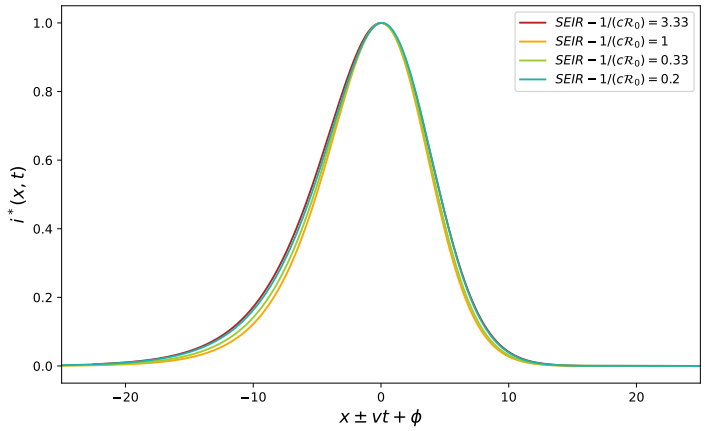


FIGURE 4.5: Scaled infected concentration $i^*(x,t)$ in the SEIR model as a function of $x \pm v^*t + \phi$ for multiple values of $1/c$ in a single figure while $\mathcal{R}_0 = 2$. v^* is the dimensionless velocity as described by Equation (3.42). ϕ is added such that the peak of the curve is at 0.

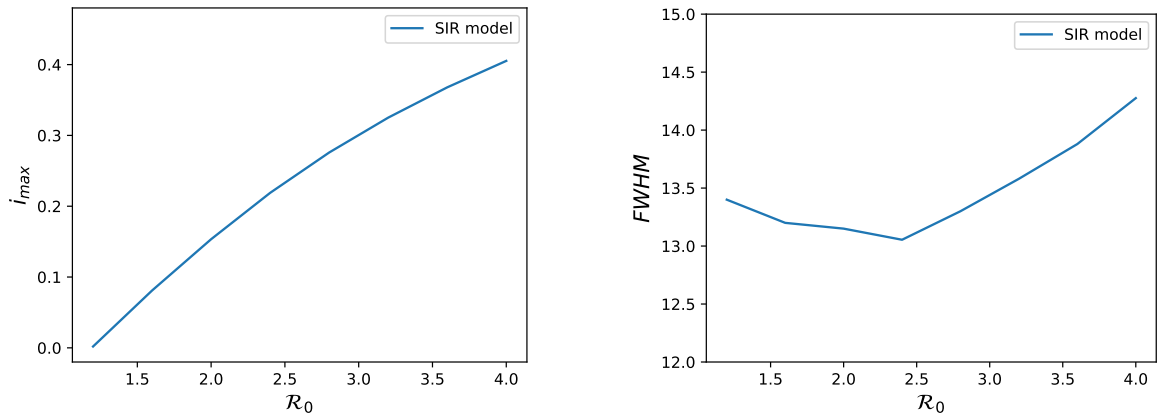
size of the epidemic r_∞. It represents the fraction of the population that has had the disease when the epidemic dies out eventually. This value only depends on $\mathcal{R}_0$ in both the SIR and SEIR models. We can calculate it now via:

$$r_\infty(x) = \int_0^\infty i(x,t)dt \tag{4.1}$$

We can think of $r_\infty(x = x')$ as the area under $i(x',t)$ for $t \in [0,\infty)$. For both models, large values of $\mathcal{R}_0$ result in higher maximum values of $i(x,t)$ and the shape becomes more asymmetric.

For the SIR model, Figure 4.6(a) shows $i_{\max}$ as a function of $\mathcal{R}_0$. $i_{\max}$ increases for larger $\mathcal{R}_0$. Plotting the results in a log-log plot gives no constant slope. The full width of the peaks at half the maximum, the FWHM, is shown in Figure 4.6(b). Although the FWHM remains approximately equal for all values $\mathcal{R}_0$, it shows an unexpected kink in the curve around $\mathcal{R}_0 \approx$

2.5.



(a) SIR model - i_max

(b) SIR model - FWHM

FIGURE 4.6: Results on the maximum value of the $i(x,t)$ curve, $i_{\max}$ and the full width at half maximum, FWHM, for the SIR model. Results are plotted for different values of $\mathcal{R}_0$. Initially, a fraction $i_0 = 0.01$ is initialized in a single subpopulation.

For the SEIR model, Figure 4.7 shows $i_{\max}$ as a function of $\mathcal{R}_0$ and the ratio of the latent period over the infectious period, $1/(c\mathcal{R}_0)$. As expected, for relatively large latent periods ($1/(c\mathcal{R}_0)$ is large), the peak intensity is low. If the latent period is relatively short, then $\mathcal{R}_0$ has an influence on $i_{\max}$ similar as in Figure 4.6(a).

For numerical evaluations with equal $\mathcal{R}_0$ but different latent periods $1/c$, r_∞, as calculated by Equation (4.1) remains constant. For a lower value of i_{max}, we would expect a wider peak. However, from Figure 4.5, we expect constant widths for different values of $1/(c\mathcal{R}_0)$. This can be explained by the role of the wave propagation velocity and its effect on the calculation of $r_\infty(x)$. For example, a slow, low peak could result in a larger $r_\infty(x)$ than a very fast high-intensity peak.

Figure 4.8 shows the FWHM also as a function of $\mathcal{R}_0$ and $1/(c\mathcal{R}_0)$. We expect the width for a given value of $\mathcal{R}_0$ is independent on $1/(c\mathcal{R}_0)$. However, we observe a small influence and that the FWHM is lowest for $1/(c\mathcal{R}_0)$ around unity. The reason for this is unclear.

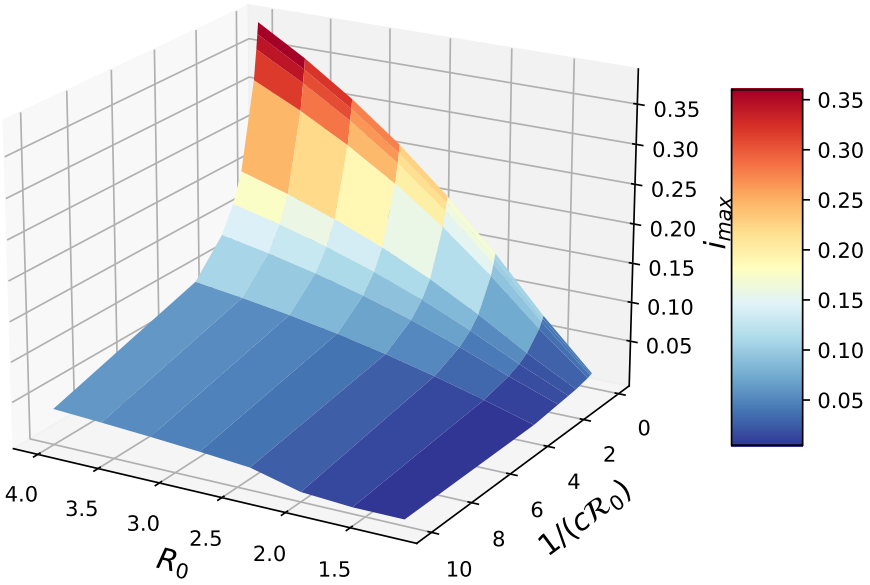


FIGURE 4.7: Results on the maximum value of the $i(x,t)$ curve, $i_{\max}$, for the SEIR model. Results are plotted for different values of $\mathcal{R}_0$ and $1/(c\mathcal{R}_0)$. Initially, a fraction $i_0 = 0.01$ is initialized in a single subpopulation.

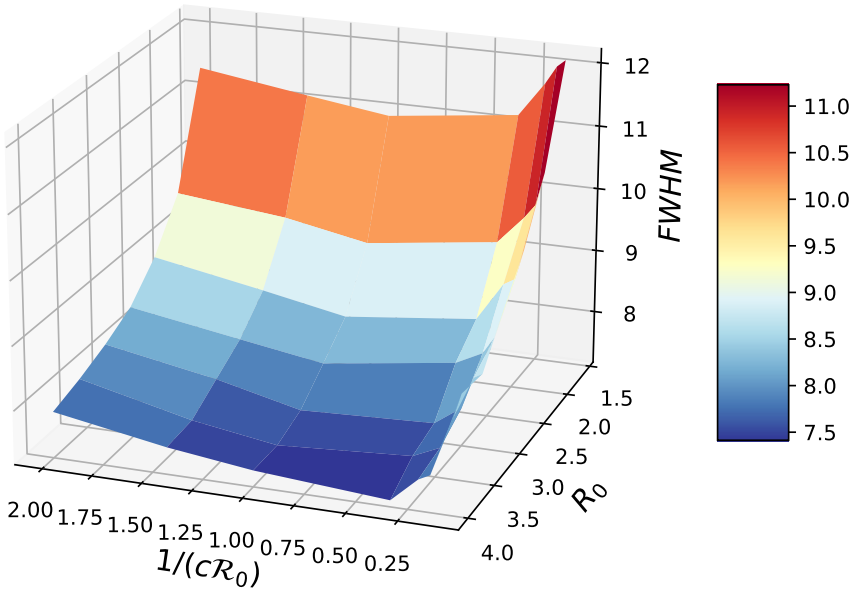


FIGURE 4.8: Results on the full width at half maximum, FWHM, of the $i(x,t)$ curve for the SEIR model. Results are plotted for different values of $\mathcal{R}_0$ and $1/(c\mathcal{R}_0)$. Initially, a fraction $i_0 = 0.01$ is initialized in a single subpopulation.

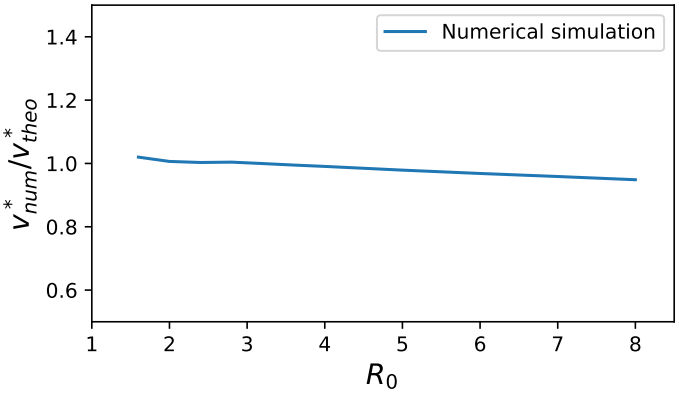


FIGURE 4.9: Numerical evaluation of wave propagation velocity v^*_{num} divided by the analytical propagation velocity from Equation (3.39) in an SIR model as a function of $\mathcal{R}_0$. The meta-population network is a circular network where mobility is only present between nearest neighbors. Initially, a fraction $i_0 = 0.01$ is initialized in a single subpopulation.

4.1.2 Propagation velocity and robustness

The propagation velocity of the epidemic wave is described by Equations (3.39) and (3.42) in terms of the model parameters. In this section, we verify both equations and show that these equations indeed correspond to the wave propagation velocity. Also, we check that the wave propagation velocity is robust to to population distribution changes, as long as $\mathcal{R}_0 > 1$.

We calculate the velocity by dividing the distance from the initial outbreak position x_0 by the time at which the wave arrives at the location t_a:

$$v^*_{num} = \frac{|x - x_0|}{t_a}. \tag{4.2}$$

We define the arrival time t_a as when the infection concentration has reached a value of $i(x,t) = 0.001$. This value is chosen arbitrarily. We could have also used the maximum value of the wave i_{max} instead of 0.01.

For the SIR model, the numerical results on the dimensionless velocity, v^*_{num}, and the theoretical results from (3.39), denoted by v^*_{theo}, correspond very well, see Figure 4.9. In this figure, we show the ratio v^*_{num}/v^*_{theo}. There is a small downward trend in this ratio. For systems with higher velocity, see Figure 3.6, we have decreased the numerical step size to improve the result.

Similarly, for the SEIR model, we show the ratio of the numerical dimensionless velocity over the theoretical numerical velocity from Equation (3.42) in Figure 4.10. The numerical velocities correspond very well. Again, for high propagation velocities, see Figure 3.6, the numerical velocity is a little lower. We reduce the numerical step size, which leads to a smaller error.

So far, we have seen that travelings waves with constant shape and velocity appear in the SIR and SEIR models if $\mathcal{R}_0 > 1$. Model parameters change the shape of $i(x,t)$. We also explored how a different population distribution influences the waves. We observe traveling waves for $\mathcal{R}_0$, with velocities predicted by Equations 3.39 and 3.42. We leave further analysis on the effect on wave properties by varying the population distribution to future work. In the next section, we discuss networks with more nearest neighbors.

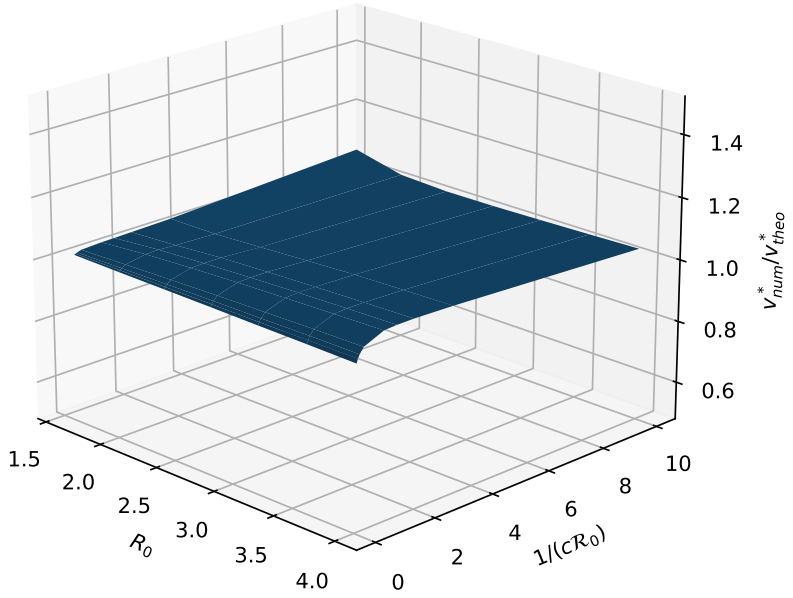


FIGURE 4.10: Numerical evaluation of wave propagation velocity v^*_{num} divided by the analytical propagation velocity from Equation (3.42) in an SEIR model as function of its model parameters, $\mathcal{R}_0$ and c. The meta-population network is a circular network where mobility is only present between nearest neighbors. Initially, a fraction $i_0 = 0.01$ is initialized in a single subpopulation.

4.2 Two dimensional network

In this section, we consider a more complex but still structured metapopulation network. We show that for a more complex network, we still see a diffusive broadening initially, which transitions into traveling waves of constant velocity and shape for $\mathcal{R}_0 > 1$. We expect that this behavior also can be observed in more complex networks with even more connected neighbors.

So far, we have considered a metapopulation as depicted in Figure 3.1. This particular network allowed us to rewrite $s_x(t)$, $e_x(t)$, $i_x(t)$, and $r_x(t)$ into $s(x,t)$, $(e(x,t))$, $i(x,t)$, and $r(x,t)$. In the $s_x(t)$ notation, x refers to the subpopulation number, see Section 2.2.1. In the $s(x,t)$ notation, x is a spatial coordinate. For a network similar as in Figure 4.11, we could rewrite all concentrations in terms of two spatial coordinates, for example., $s_x(t) \rightarrow s(x,y,t)$. The network in Figure 4.11 resembles a two-dimensional grid.

We numerically evaluate our model on a network using Equations (2.41), (2.42), and (2.43) for the SIR model and Equations (2.44), (2.45), (2.46), and (2.47) for the SEIR model. We use four nearest neighbors (network inspired by a Von-Neumann lattice) or eight nearest neighbors (network inspired by a Moore lattice); see Figure 4.11. We use a homogeneous population distribution. We use a total mobile fraction equal to $m \equiv \sum_{y(x)} m_{xy}$, where $y(x)$ represents subpopulations y connected to x, in other words, the nearest neighbors of x. For four neighbors, this results in $m_{xy} = m/4$, while for eight nearest neighbors, this results in $m_{xy} = m/8$.

We expect to see diffusive behavior, traveling waves for $\mathcal{R}_0 > 1$, and collisions for periodic boundaries or multiple outbreak locations like we saw for our one-dimensional networks in the

previous sections. Our results confirm these expectations. Namely, we see the concentration of infectious individuals grows diffusively shortly after an outbreak. Then, it turns into a circular expanding traveling wave. The exact shape of the wavefront depends on how many nearest neighbors we include; see Figure 4.12.

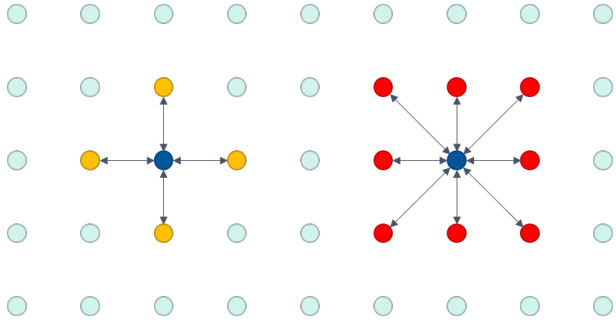


FIGURE 4.11: A metapopulation network inspired by a lattice with on the left a location with its four nearest neighbors in yellow and on the right a location with its eight nearest neighbors in red.

When comparing the top and bottom row of Figure 4.12, we see that the shape is more circular for eight nearest neighbors than for four neighbors. The wavefront velocity is not constant across the traveling wavefront as it depends on the number of infectious neighbor populations.

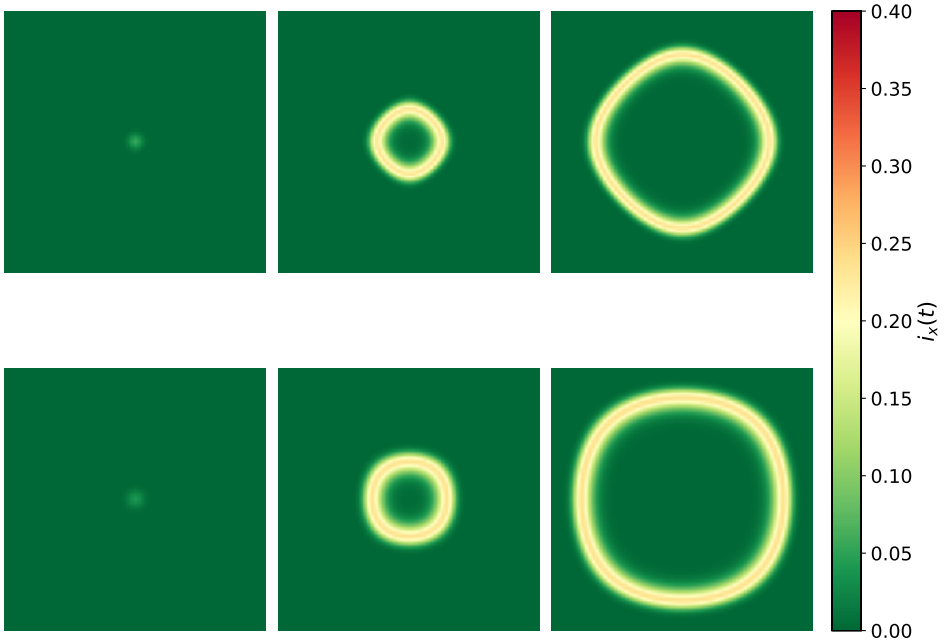


FIGURE 4.12: Two numerical evaluations of the infected concentration $i_x(t)$ in an SIR model with $\mathcal{R}_0 = 2$ on a 70x70 grid with periodic boundaries and a single outbreak location. Initially in this outbreak location x_0, we set $s_{x_0}(t) = 0.99$, $i_{x_0}(t) = 0.01$ and $r_{x_0}(t) = 0$. The infected concentration at all locations is plotted for different moments in time. From left to right, $t = 50$, $t = 150$, and $t = 300$. Each location has four nearest neighbors in the top row, while in the bottom row, each location has eight nearest neighbors, see Figure 4.11.

Compared to a one-dimensional network, the wave propagation velocity is lower for a two-dimensional network. To understand this, we need to realize that the force that drives transmission via neighboring subpopulations is proportional to m_{xy} for each neighbor, see Equation (2.41) and (2.44). We kept the total mobile fraction m of a population constant; $m \equiv \sum_{y(x)} m_{xy}$, where $y(x)$ represents subpopulations y that are connected to x, in other words, the nearest neighbors of x.

In a one-dimensional network, we have two neighbors, so m_{xy} is set to $m/2$. In a two-dimensional network, the force that drives transmission via neighboring subpopulations is proportional to $m/4$ or $m/8$, for a network with four or eight nearest neighbors, respectively. A higher force that drives transmission of the disease also means a higher wave propagation velocity.

For a localized initialization, a subpopulation is *attacked* by the epidemic via one neighbor in a one-dimensional network. The force and the velocity are proportional to $m/2$. In a two-dimensional network with four nearest neighbors, the force and velocity are proportional to $m/4$ per infectious neighbor.

For a localized initialization, we can easily see that the number of infected neighbors at the wave-front is less than er equal to two. This results in a force of infection and wave propagation velocity, which is proportional to $\leq 2m/4 = m/2$. As a result, the velocity in a 2D grid is smaller or equal to the velocity in a 1D network for a localized initial outbreak.

Multiple outbreak locations. Here, we simulate multiple outbreaks in combination with eight nearest neighbors. The outbreaks locations relatively close to each other; the distance equals $6\Delta x$, where Δx is the distance between two neighbors. We expect the separate wave-fronts to combine into a single wavefront of circular shape. In Figure 4.13, we see exactly happening what we expected.

When the traveling wavefront radius becomes much larger than the initial separation in outbreak locations, the shape of the wavefront approaches Figure 4.12. In other words, as if the epidemic had a single source instead of multiple sources.

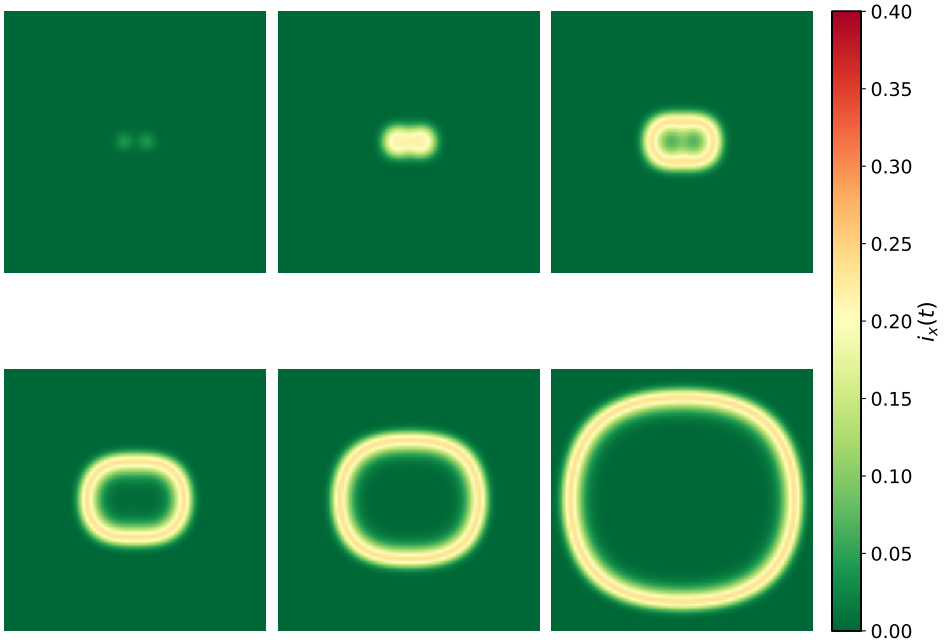


FIGURE 4.13: A numerical evaluation of the infected concentration $i_x(t)$ in an SIR model with $\mathcal{R}_0 = 2$ on a 70x70 grid with periodic boundaries and two outbreak locations. Initially in this outbreak location x_1 and x_2, we set $s_{x_1}(t) = s_{x_2}(t) = 0.99$, $i_{x_1}(t) = i_{x_2}(t) = 0.01$ and $r_{x_1}(t) = r_{x_2}(t) = 0$. The infected concentration at all locations is plotted for different moments in time. From left to right, first row, $t = 50$, $t = 80$, and $t = 110$. From left to right on the second row $t = 150$, $t = 200$, and $t = 300$. Each location has eight nearest neighbors. The wave fronts collide and merge into a single larger traveling wave front.

Chapter 5

Mobility restrictions

In this chapter, we present the article published in The Royal Society Interface, entitled: *Trade-offs between mobility restrictions and transmission of SARS-CoV-2* [9]. The article has been slightly modified to fit to the narrative and style of the thesis, some sections are rewritten or left out.

In this chapter, we address research question 2: *How to balance the trade-off between mobility restrictions and transmissions of COVID-19?*. We formalize this trade-off in terms of a single trade-off parameter. This parameter can be interpreted as the number of movements between areas that we are willing to restrict to prevent the occurrence of a *single* new infection. We consider restrictions where we divide a country into different regions. We block the mobility between different regions but allow for mobility within these regions. We use this definition to answer the two subsequent questions: *How to evaluate the effectiveness of mobility restrictions?* And *How large should a 'region' be for effective containment of the disease?*. We use an SEIR epidemic model to evaluate the impact of mobility restriction and propose heuristic methods to approximate optimal choices for mobility restrictions.

The author of this thesis is a co-author of the article. The contributions of all co-authors are as follows:

- Martijn Gösgens MSc. formalized the problem, helped develop the methodology and drafted the article;
- Teun Hendriks helped formalize the problem, helped develop the methodology, analyzed the model and drafted the article;
- Dr. ir. Marko Boon helped set up the project, helped obtain the data and critically revised the manuscript;
- Wim Steenbakkers provided the mobility data;
- Prof. dr. ir. Hans Heesterbeek contributed to the discussions and helped draft the main text;
- Prof. dr. Remco van der Hofstad contributed to the discussions and helped draft the main text;
- Prof. dr. Nelly Litvak contributed to the discussions and helped draft the main text.

5.1 Introduction

At any given moment in the outbreak of a virus-transmitted disease, there is a mix of countries and regions where the virus is temporarily under control, countries where the epidemic is decreasing, and countries where the epidemic is increasing. After an initial peak in cases, countries remain at risk for second and subsequent peaks, even if no cases are reported in the country for long periods of time. As in principle, everybody is susceptible to some degree, not reaching herd immunity (see section 2.1.4) after the initial wave of infection leaves a largely susceptible population that can sustain subsequent outbreaks [31].

Mobility between regions plays a potentially important role in increasing transmission. Still, measures aimed at restricting mobility also have a potentially large social and economic impact: mobility and economic activity are often studied as two sides of the same coin [28]. In the initial wave of infection, countries, to a large extent, impose national mobility restrictions. More regional or local mobility restrictions can perhaps achieve containment strategies for preventing subsequent infection waves. This has the advantage of reducing the social and economic burden on society yet has the risk that the restrictions may not be sufficiently effective and need to be scaled-up after all to a national level at some later point in time.

It is, however, unclear how one could gauge the effectiveness of regional restrictions based on realistic mobility patterns specific to the country, balancing trade-offs between mobility and transmission. It is also unclear how large a 'region' should be for effective containment and how different recognizable regions (for example, administrative regions such as provinces, large cities, or postal code regions).

In the remainder of this chapter, we provide a framework to evaluate regional strategies' effectiveness to restrict mobility, allowing for a range of choices of how regions are characterized, using The Netherlands as a case study. This is essential to determine the scale at which interventions can be effectively imposed or lifted, and addresses a range of key modeling questions for COVID-19 and future pandemic outbreaks [68].

We base the framework on actual mobility patterns in The Netherlands. We distinguish between extreme situations where the infection is distributed evenly between areas, and situations where the infection is highly concentrated in a restricted number of areas, for example, as a result of a super spreading event. We show that regions defined based on mobility patterns provide better strategies than regions based on administrative characteristics and that focusing on administrative regions, therefore, leads to sub-optimal strategies. We also quantify and explore the non-linear relation between mobility and outbreak size for a range of choices of the trade-off between mobility and transmission.

5.2 Methodology

In this section, we describe the overview of our approach. We consider restrictions where we divide a country into different regions. We block the mobility between different regions but allow for mobility within these regions. This chapter aims to devise regions that allow for as much mobility as possible yet restrict infections as much as possible. For this, we need to strike a careful balance between mobility and infections, which we formalize in terms of a trade-off parameter that policymakers need to impose.

This section is organized as follows. We start in Section 5.2.1 by specifying the kind of mobility strategies that we consider in this work. As we want to balance infections and mobility, we describe how we estimate those in Section 5.2.2 and Section 5.2.3, respectively. Next, in Section

5.2.4, we introduce a way to quantify mobility restriction strategies' performance by formulating the trade-off between mobility and infection spread as an explicit optimization problem over the various choices of regions described in Section 5.2.1.

The optimization problem that formalizes our trade-off between mobility and infection containment is inspired by community-detection algorithms and is complex to solve explicitly. In Section 5.2.4, we also describe how to rigorously and heuristically analyze such problems. In Section 5.2.5, we provide heuristics that generate strategies with high performance. We close this section by describing how we can evaluate the various divisions in regions in Section 5.2.6.

5.2.1 Strategies for mobility restrictions

Within the meta-population framework (see Section 2.2.1), a region is represented by a subset $D \subseteq \mathcal{A}$. For locations $x, y \in D$, the number of daily travels between x and y, M_{xy}, and the scaled mobility $m_{xy} = M_{xy}/N_y$ is left unchanged, while for $z \notin D$, $M_{xz} = 0$ is set to zero. Finally, a division is represented by a partition $\mathcal{D} = \{D_1, \ldots, D_{|\mathcal{D}|}\}$. In our use case, we consider Dutch municipalities to be sub-populations.

Many administrative divisions that might serve as examples for regions $\mathcal{D}$ already exist. For example, we could use The Netherlands its divisions of its municipalities into 12 provinces or 25 so-called security regions[1]. An advantage of using such divisions is that they are already known so that it may be easier to communicate and thus enforce mobility restrictions based on them to the broad public. However, a disadvantage is that these divisions have been historically determined by decisions of governance. Thus their borders do not necessarily effectively reflect the actual movement of people throughout the country. As a result, mobility restrictions based on them may not be the most effective.

An illustrative example is the province of Flevoland (equal to the security region Flevoland). The city of Almere, the most populous city of Flevoland, lies close to the city of Amsterdam, where many of the city of Almere's citizens work. Our mobility data shows that more than 90% of the mobility leaving Almere also leaves the province of Flevoland, as can be seen in Figure 5.5. Therefore, choosing the division into provinces or security regions would disproportionately affect Almere's people in terms of mobility. In Section 5.2.5, we provide a method to obtain divisions based on the mobility data by applying community detection methods. In Figure 5.5, we see that this method's divisions do consistently place Almere in the same region as Amsterdam.

Another disadvantage of using existing administrative divisions is that these are, by their very definition, inflexible and can hence not be tailored to the specific epidemiological situation. In Section 5.2.5, we provide a method to obtain divisions that do take epidemiological information into account.

5.2.2 Estimating infections

Infections are estimated with a SEIR remote-force-of-infection model, see Section 2.2.3. The meta-population structure is based on The Netherlands and its municipalities.

We distinguish between *Tested* and *Untested* individuals. When individuals become infectious, a fraction a gets tested, while the remainder $(1 - a)$ remains undetected. This distinction allows

[1] In The Netherlands, a safety region is an region in which various authorities and services work together to carry out tasks in the field of fire departments, disaster and crisis management, medical assistance, public order and safety.

us to initialize based on historical reported numbers of new infections easily. An important assumption is that tested individuals do not transmit the disease to other locations. In other words, they remain within their municipalities. This also means when $a = 1$, all infectious individuals are known, and the disease does not spread to other locations.

This means we can rewrite infection rates between different locations from Equation 2.40 such that only a fraction $(1 - a)$ of the infectious individuals transmits the disease to other locations:

$$\begin{cases} \beta_{xy} = \beta_l & \text{if } x = y, \\ \beta_{xy} = (1-a)(m_{xy} + m_{yx})\beta_m & \text{if } x \neq y, \end{cases} \tag{5.1}$$

It is possible to calculate the actual number of people in the compartments via $S_x = s_x N_x$, $E_x = e_x N_x$, etc., where N_x is the population size of location x. We present the parameter estimates in Table 5.1. We estimate $\beta_l = c_l \varepsilon$ and $\beta_m = c_m \varepsilon$, where c_l and c_m are the local contact rate and the non-local contact rate and ε is the transmission probability. In Section 2.1, we explained that the infection rate β could be written as a product of c and ε.

Based on the literature, we estimate that the average number of contacts a person has per day is 13.85 in The Netherlands [69], but the transmission probability is unknown. We assign ε a value such that the $\mathcal{R}_0$ we calculate corresponds to the $\mathcal{R}_0$ of COVID-19.

We assume that the fraction of local contacts equals $p = 0.5$, which means that $c_l = pc$ and the remainder of contacts are made when individuals travel. We calculate the $\mathcal{R}_0$ using the Next Generation Matrix (see Section 2.1.4). For $p = 0.5$, we want $\mathcal{R}_{eff} = 1.25$. This is not the actual $\mathcal{R}_0$ of SARS-CoV-2, which is around 2.5 [46], but this 1.25 considers the efficacy of prevention measures such as social distancing and wearing masks. By reducing mobility, we can then reduce $\mathcal{R}_{eff}$ to values lower than 1. This leads to effective containment of the epidemic. For more details on how these parameters are estimated, we refer to Appendix F.

For $a = p = 0.5$, we set $\mathcal{R}_{eff} = 1.25$. When we vary the fraction of people who get tested a and the fraction of local contacts p, we obtain different reproduction numbers. Figure 5.1 gives insights into how the different combinations of a and p influence the reproduction number. For example, if one would find almost all infectious individuals by testing $(a \approx 1)$, in combination with people having 50% of their contacts locally $(p = 0.5)$, the reproduction number is below 1 without mobility restrictions. We also see that if we do not test enough, the reproduction number is likely to be larger than 1 and additional mobility restrictions are necessary for containment. For a list of parameters, we refer to Table 5.1.

In combination with initial conditions, we can simulate how the concentration evolves. We numerically calculate $s_x(t)$, $e_x(t)$, $i_x(t)$, and $r_x(t)$ via a forward Euler scheme, that needs initial conditions $s_x(t_0)$, $e_x(t_0)$, $i_x(t_0)$, and $r_x(t_0)$.

$$f_x(t + \Delta t) = f_x(t) + \Delta t \frac{df_x(t)}{dt}, \tag{5.2}$$

where $\delta t = 1$ day. Using this method, we can estimate the number of infections for any number of days after initialization. Let $G(\mathcal{D}, T)$ be the number of new infections within a certain total time T, which depends on the mobility restriction by division $\mathcal{D}$. This can be calculated using the total number of susceptible individuals:

$$G(\mathcal{D}, T) = \sum_{x \in \mathcal{A}} N_x s_x(t_0) - \sum_{x \in \mathcal{A}} N_x s_x(t_0 + T) \tag{5.3}$$

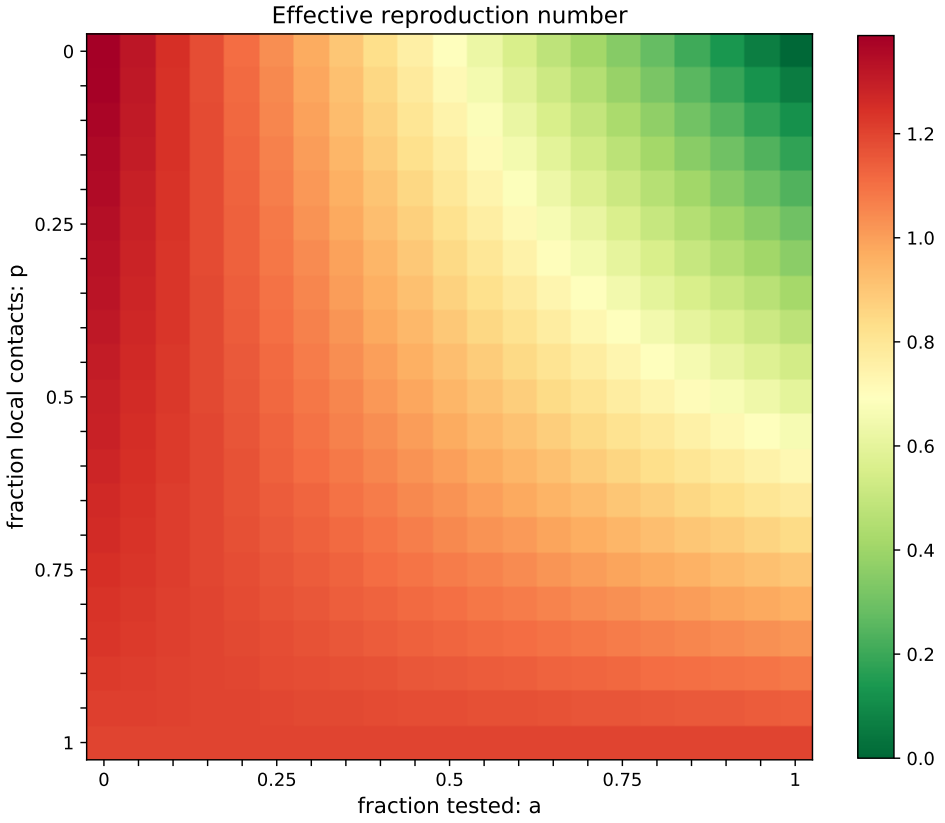


FIGURE 5.1: Effective reproduction number $\mathcal{R}_{eff}$ as a function of the fraction of individuals who are tested a and the fraction of contacts each individual has within its own municipality p. We use $1/\gamma = 5$ days, $c = 13.85$ contacts per day, and a transmission probability of $\varepsilon = 0.0238$.

Name		Value	Source
Fraction tested	a	1/15	estimated
Fraction local contacts	p	1/2	estimated
Infectious period	$\frac{1}{\gamma}$	5 days	Deng et al [71]
Latent period	$\frac{1}{\nu}$	4 days	Shorter than incubation period [41, 72]
Basic reproduction number	$\mathcal{R}_0$	2.5	Li et al. [46]
Prevention measure efficacy		50%	estimated
Effective reproduction number (with prevention measures)	$\mathcal{R}_{\text{eff}}$	1.25	50% of $\mathcal{R}_0$
Transmission probability per contact with infectious person	ε	0.0238	NGM and $\mathcal{R}_{\text{eff}} = 1.25$
Average contact rate (unique persons)	c	13.85	Mossung et al. [69]
Transmission rate via local contacts	β_{l}	0.165	$\beta_{\text{loc}} = c_{\text{loc}}\varepsilon$
Transmission rate via mobility related contacts	β_{m}	0.141	$\beta_{\text{mob}} = c_{\text{mob}}\varepsilon$

TABLE 5.1: Values of the parameters in our epidemic model.

Here, $s_x(t)$ is the fraction of susceptible individuals in subpopulation x, which has population size N_x. We use that $s_x(t)$ is a decreasing function (see Section 2.1.4) and individuals only leave the susceptible compartment because they are infected.

We consider three different initialization methods. One initialization is based on actual historic infection numbers reported by the Dutch authorities, referred to as *historic initialization* [70]. The other two demonstrate scenarios where initially the infectious are either evenly distributed across the country or where they are concentrated in a single municipality. All three are described in more detail in Appendix F.

To visualize how concentrated the epidemic is, we introduce a concentration measure. This measure compares the population distribution of the different locations with the infection distribution. The measure is equal to zero if the distributions are equal, which is the case after an evenly distributed initialization. The measure is close to unity if it is extremely concentrated. For historic initializations, it fluctuates between these two values. We present the results in Figure 5.4. For more details on how this concentration measure is calculated, we refer to Appendix F.

5.2.3 Estimating mobility

We estimate mobility in the model using the Mezuro Mobility data. The data is collected in 2019 between the 1st of March and the 14th of March. It is aggregated to a daily average. We have a daily average of how many people move between their base location and destination locations for each municipality. We do not distinguish between working week and weekend days. The period between March 1 and March 14 is not typical. It contains *Carnaval*, an event in the southern provinces of The Netherlands, which attracts people from all regions of the country. Therefore, we expect the southern municipalities to attract more visitors in our dataset compared to average weeks.

The mobility data set contains M_{xy} for x and y municipalities of The Netherlands, which is the average number of daily visitors who live at y and visit x. This daily average is equal to M_{xy} from Section 2.2.3. While it is not an ideal data set, we still use it as a proxy for actual mobility patterns in The Netherlands. It serves as input for simulating epidemics in The Netherlands. We can also scale the daily mobility to m_{xy} by using the population size at y, $m_{xy} = M_{xy}/N_y$.

Some interesting insights show the strong heterogeneity in population size distribution and the degree distribution. Also, the average shortest path between any two municipalities is less than 2, making it difficult to contain an epidemic.

The nodes in the network are Dutch municipalities. In 2019, The Netherlands had 380 municipalities. The population sizes of different municipalities vary between a little under 1000 inhabitants up to around 900.000 (Municipality of Amsterdam) with an average value of around 45.000 inhabitants, see Figure 5.2(a).

The total degree of the nodes (the number of other nodes that have a direct connection, or in other words, the number of nearest neighbors) is in the range 1 to 352 with an average of just above 59 (see Figure 5.2(b)). We observe that some nodes have almost the maximum degree, 380, so they are connected to almost all other nodes. There appears to be a linear relationship between the population size and the degree of the municipalities for smaller municipalities, see Figure 5.2(c).

The weight of edges (number of daily travels between two municipalities) lies in the range of 36 to just over 48.000. This results in a total number of daily visitors for each node between 85 and just over 540.000 with an average of 28.000 daily visitors, see Figure 5.2(d).

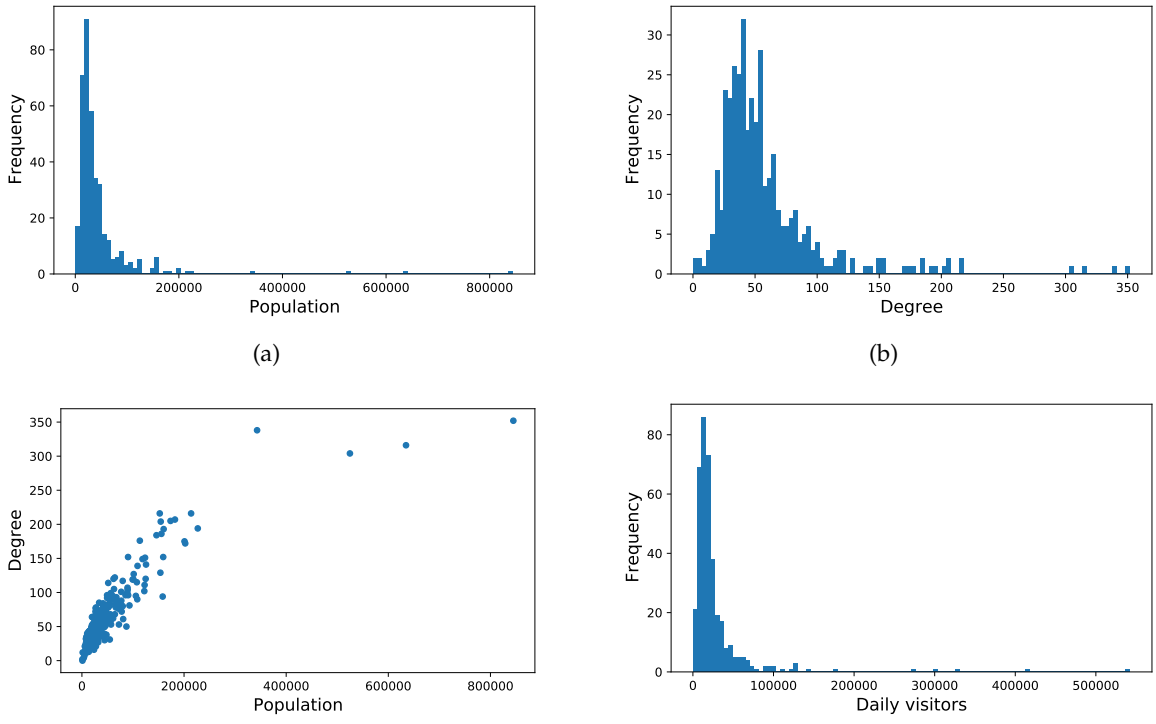


FIGURE 5.2: Descriptive data analysis on the mobility data set of The Netherlands. Municipalities are nodes in the network. Weighted edges represent the average number of daily travels. (a): Histogram of population size distribution. (b): Histogram of degree distribution. The degree equals the number of connected municipalities a certain municipality has. (c): Degree versus the population size (d): Histogram of number of daily visitors.

Secondly, based on the division in regions $\mathcal{D}$, some mobility between municipalities is not allowed, because this is how we defined mobility restrictions. For these connections, $M_{xy} = 0$.

Without restrictions, the total daily mobility equals

$$M_{\text{tot}} = \sum_{x,y \in \mathcal{A}} M_{xy}. \tag{5.4}$$

With mobility restrictions, i.e., a division into regions, the allowed mobility is

$$M_{\text{allowed}}(\mathcal{D}) = \sum_{D \in \mathcal{D}} \sum_{x,y \in \mathcal{D}} M_{xy}. \tag{5.5}$$

From this, we can calculate the percentage allowed mobility via:

$$\textit{percentage allowed mobility} = \frac{M_{\text{allowed}}(\mathcal{D})}{M_{\text{tot}}} \times 100\%. \tag{5.6}$$

Let $\mathcal{M}(\mathcal{D}, T)$ be the actual allowed mobility within a total time of T. This can be calculated via multiplication:

$$\mathcal{M}(\mathcal{D}, T) = M_{\text{allowed}}(\mathcal{D})T. \tag{5.7}$$

5.2.4 Objective

On the one hand, freedom of movement has both economic and intrinsic value. On the other hand, it also facilitates the spread of the disease. In essence, the problem for control is to find a trade-off between mobility and infection containment, given the epidemiological characteristics and normal mobility patterns. In this section, we provide a way to *formalize* this trade-off to allow its characterization.

Trade-off parameter. To formalize the trade-off between public health and mobility, we introduce a trade-off parameter α. This parameter can be interpreted as the number of movements between areas that we are willing to restrict to prevent the occurrence of a *single* new infection. A higher value for α thus favors more severe restrictions. This trade-off parameter reflects societal values and is a political choice that politicians or policy makers should make. Therefore, we refrain from advising about a specific suitable value but instead provide a method that advises a strategy given a choice for α.

Time horizon. Suppose that a certain set of mobility restrictions is in place. We consider some time horizon of a total number of T days representing the number of days that the restriction will be in place and count the number of infections and movements before this horizon. This time horizon should thus not be too long, as it coincides with the duration that restrictions are in place, and sometimes one needs to quickly respond to changes in the infection spread quickly. Due to the delay between transmission and becoming infectious(latent period), the time horizon should also not be too short, as otherwise, the effect of the imposed restrictions cannot reasonably be observed. In this work, we use a time horizon of $T = 30$ days as an example.

At the end of this time horizon, a new division into regions may be chosen based on the status of the epidemic at that moment in time, thus allowing for a dynamic update of the strategy of mobility restrictions.

Objective function. The above considerations lead to the objective function. In this function, we balance the allowed mobility versus the new infections within a time horizon T. We want

to maximize the objective function:

$$Q_{\gamma,T}(\mathcal{D}) = \mathcal{M}(\mathcal{D}, T) - \alpha G(\mathcal{D}, T), \tag{5.8}$$

where $\mathcal{M}(\mathcal{D}, T)$ and $G(\mathcal{D}, T)$ represent the total allowed mobility (5.7) and the total number of new infections (5.3), respectively, that occur before the time horizon T, given a division $\mathcal{D}$. Note that in this formulation the current status of the epidemic, i.e., the number of infectious and susceptible people at the start of the period T, is included because the value of $G(\mathcal{D}, T)$ depends on that initial status.

The objective function (5.8) should be interpreted as follows: given divisions $\mathcal{D}$ and $\mathcal{D}'$, if $Q_{\alpha,T}(\mathcal{D}) > Q_{\alpha,T}(\mathcal{D}')$, then $\mathcal{D}$ is preferred over $\mathcal{D}'$ with respect to the trade-off given by α. Hence, the objective function establishes an ordering among the divisions. This study aims to provide a methodology that, given α and T, finds a division $\mathcal{D}$ with a (locally) optimal $Q_{\alpha,T}(\mathcal{D})$.

Optimization The quantification of $Q_{\alpha,T}(\mathcal{D})$ in Equation (5.8) is computationally expensive. Firstly, because it involves computing the number of infections at the end of the time horizon and secondly, because the number of possible divisions $\mathcal{D}$ is enormous. For a set of size n, $\mathcal{A} = [1, 2, .., n]$, the number of divisions (subsets) can be calculated via:

$$\sum_{i=1}^{n} i! \binom{n}{i} = \sum_{i=1}^{n} i! \frac{n!}{i!(n-i)!}. \tag{5.9}$$

For $n = 380$, the number of Dutch municipalities, the number of possible divisions is enormous and therefore it becomes not feasible to find the global optimum. Therefore, we will resort to heuristic optimization methods. We describe these methods in Appendix G.

The idea of our approach is to optimize existing divisions and new divisions (see next section). For example, combining multiple provinces into larger regions might not strongly influence the spread of the disease, while it could allow for a lot more mobility. This approach is much faster because the number of regions in a division (12 provinces) is usually much smaller than starting with all the individual municipalities ($n = 380$), which drastically reduces the number of divisions, see Equation (5.9).

5.2.5 New divisions

We mentioned before that administrative divisions do not reflect the movement of people throughout the country, nor do they take into account the current state of the epidemic. Here, we present *Mobility regions* and *Adaptive mobility regions*.

Mobility regions. The mobility region method is based on the Louvain algorithm, a community detection algorithm. This algorithm finds communities by optimizing a quantity called *modularity*. It combines sub-population, which are densely connected into communities.

A single resolution parameter determines the granularity or sizes of these communities. This method is described in more detail in Appendix G. An example outcome is shown in Figure 5.5. For example, the city of Almere is now in a similar region as the city of Amsterdam in the same region. Also, we show two examples in Figure G.1.

Adaptive mobility regions. We use the Louvain algorithm, but instead of optimizing the modularity, we optimize a modified function that takes into account the force of infection. This

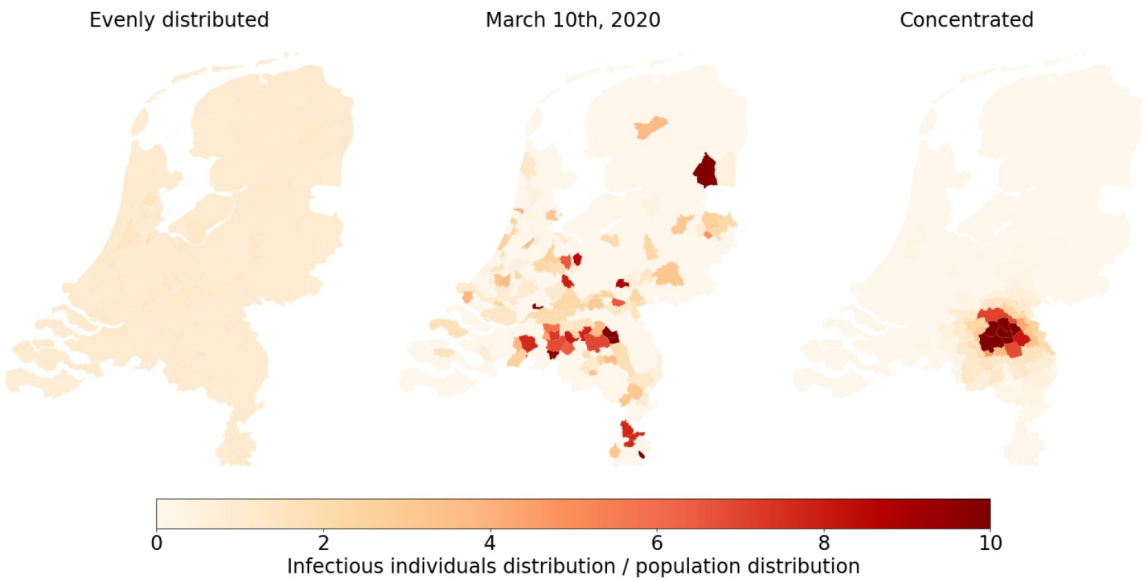


FIGURE 5.3: Ratio of infectious individuals distributions over population distribution. The model is initialized with either evenly distributed, historical, or concentrated distributions of active infections. For the evenly distributed an concentrated initializations, a total of 1000 people are initially exposed and 10 days are simulated. The infectious individuals distribution equals $\rho_x^{\text{inf}}(t) = \frac{i_x(t)N_x}{\sum_{x\in\mathcal{A}} i_x(t)N_x}$. The population distribution equals $\rho_x^{\text{pop}} = \frac{N_x}{\sum_{x\in\mathcal{A}} N_x}$.

method is described in more detail in Appendix G. Also, we show two examples in Figure G.2.

Our implementation results in large communities when the number of infectious individuals is low in all sub-populations and small communities when the number of infectious individuals is high. Again, a single resolution parameter controls the number of different communities the method obtains. An example is shown in Figure 5.6, where you can clearly see the smaller regions (communities) where the number of infections per capita is high.

5.2.6 Evaluation of mobility restrictions

Given values for the trade-off parameter α and the time horizon T, we can compute the value $Q_{\alpha,T}(\mathcal{D})$ to assess the quality of some newly obtained division $\mathcal{D}$. However, by itself $Q_{\alpha,T}(\mathcal{D})$ has no clear interpretation. We can obtain insight by comparing the quality of $\mathcal{D}$ to the quality of existing alternative divisions. These existing divisions are referred to as benchmark divisions. There exist administrative choices to divide the country into regions, such as provinces and security regions. We consider the following benchmark divisions for the situation of a widespread outbreak:

- No restrictions (minimum restrictions);
- Disallow movement between provinces;
- Disallow movement between security regions;
- Disallow movement between municipalities (maximum restrictions).

Note that minimum restrictions maximize mobility and infections, while these are both minimized by maximum restrictions. For a concentrated initialization in a single municipality, we also consider the following alternative benchmark divisions:

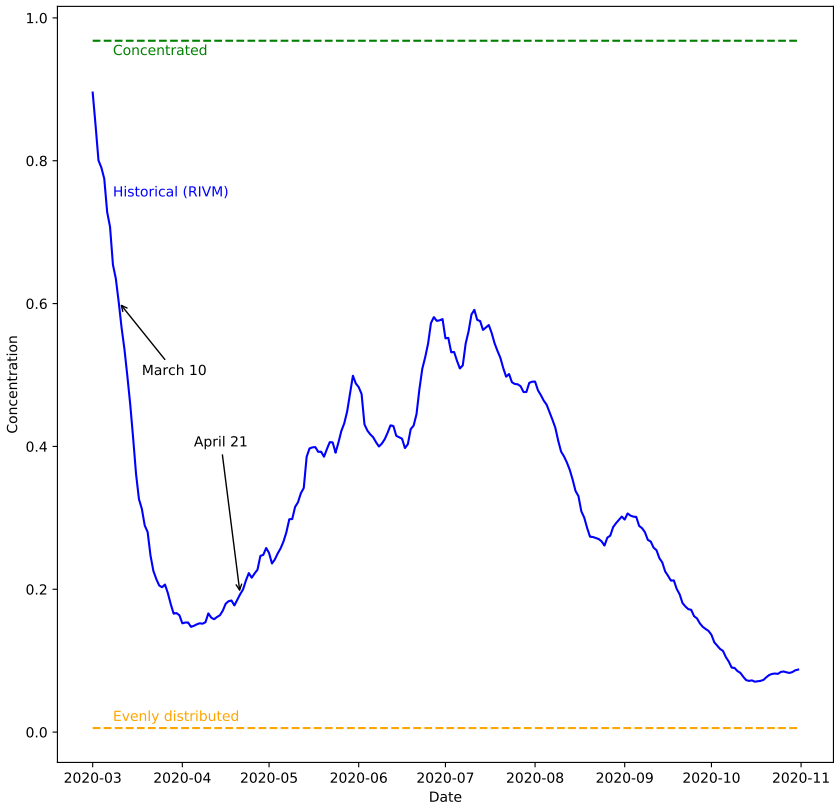


FIGURE 5.4: The concentration values corresponding to distributions of infections according to the RIVM data over time, see Appendix F for calculation of concentration values. The dotted lines show the concentration levels of concentrated and evenly distributed initializations, see Appendix F. Data obtained from https://data.rivm.nl/covid-19/.

- Isolation of the municipality;
- Isolation of the security region of this municipality;
- Isolation of the province of this municipality.

Given a value for the trade-off parameter, we can assess whether the division that is obtained by optimizing the objective for this trade-off value indeed outperforms the benchmark divisions.

Finally, given a set of divisions $\mathcal{D}_1, \ldots, \mathcal{D}_k$ and a time horizon T, we can plot (5.6) and (5.3) for each division, resulting in plots such as in Figures 5.10(a), 5.10(b), 5.11(a), and 5.11(b). Note that we scale (5.3) with the total population size in our results.

In these plots, if a division is plotted to the right of a benchmark division, it is more favorable in mobility, while a lower vertical position indicates fewer infections. if a division is such that both are the case, we can say that a division *dominates* the benchmark division: for any value of the trade-off parameter, it is favored over the benchmark division. Note that it is not possible to dominate the minimum and maximum restrictions benchmarks since they achieve maximum mobility and minimum infections, respectively.

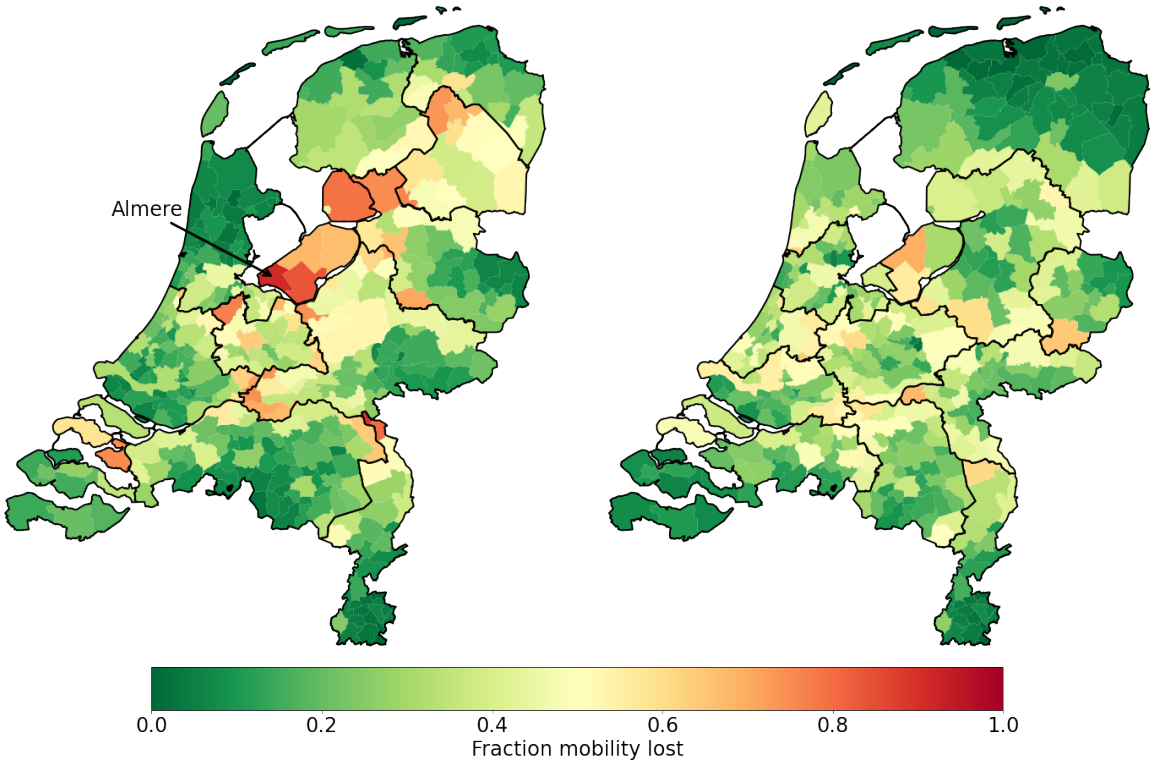


FIGURE 5.5: Each municipality is colored based on the percentage of mobility to destinations outside its region. Left: regions are provinces. Right: regions are mobility regions found by modularity optimization with resolution parameter= 2, chosen such that both divisions consist of 12 regions.

Given a set of divisions $\mathcal{D}_1, \ldots, \mathcal{D}_k$ and a choice for the trade-off parameter and the time horizon, we can select the division with the highest objective value. Obviously, the resulting division is not guaranteed to be the global optimizer of the objective, as this would require comparing an enormous number of divisions. Therefore, the quality of such choice depends on the quality of the candidates $\mathcal{D}_1, \ldots, \mathcal{D}_k$.

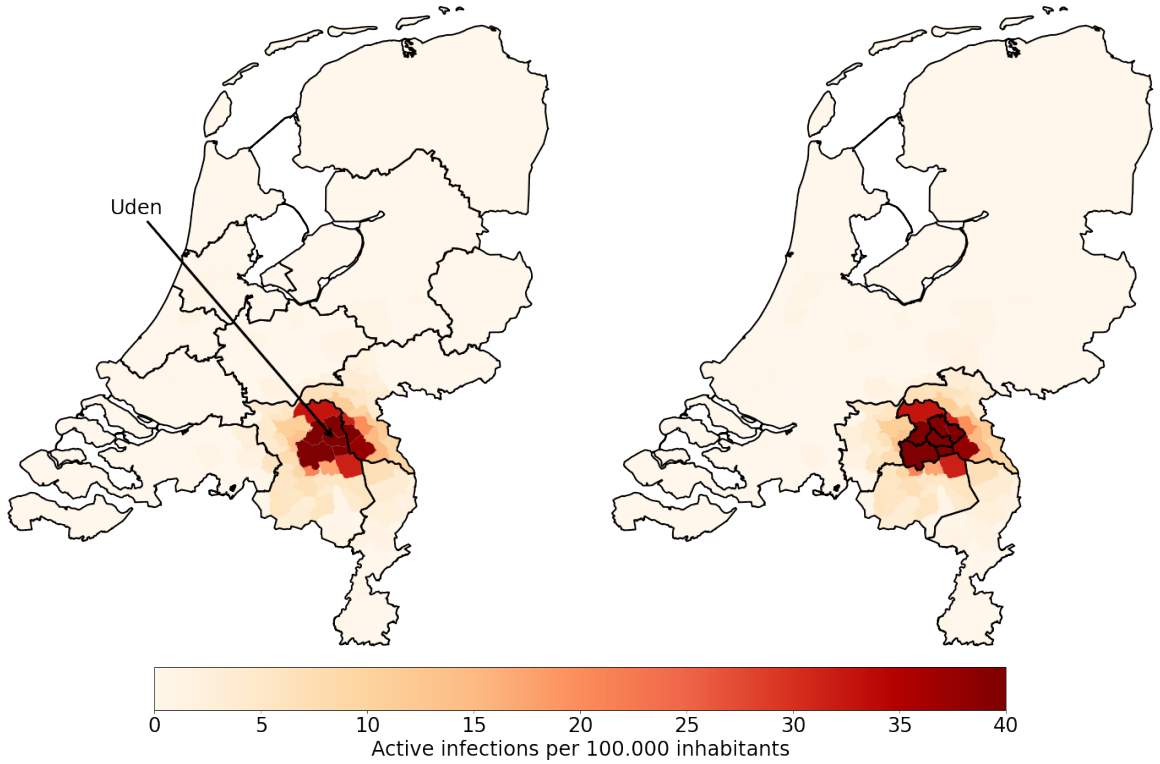


FIGURE 5.6: The mobility regions on the left versus the adaptive mobility regions on the right. The colors denote the number of active infections per capita for the concentrated initialization in the municipality of Uden.

5.3 Stochasticity and super spreaders

Before we evaluate mobility restrictions, we analyze the effects of stochasticity and, in particular, the effect of super spreaders. We expect stochastic effects to play a significant role when the number of infections is low, and the system is small [11]. We see in stochastic simulations that indeed, only for small numbers of infections, stochastic effects are significant. For larger systems, the stochastic results are very close to the deterministic mean-field results, see Appendix H.

The reproduction number tells us something about the average new cases caused by an infectious individual, see Figure 2.8. However, the reproduction number of certain individuals can be very different. We refer to these individuals with high reproduction numbers as super spreaders, see Figure 5.7. For COVID-19, it appears that the reproduction number is on average 2.5 [46], but it varies a lot between individuals. Recent work found that over 80 percent of the new cases are caused by approximately 10 percent of the current infectious individuals [30, 73], see Figure 5.8.

We analyze the effects of super spreaders via an SEI_2R model, where we distinguish two different types of infected individuals; non-contagious individuals and super spreaders. Non-contagious individuals are most common but do not transmit the disease, we set their reproduction number equal to zero, $R_{eff}^{\text{normal}} = 0$. The probability of becoming a super spreader is set to $p = 0.1$. Then, the infection reproduction number of the super spreaders, $\mathcal{R}_{eff}^{\text{super}}$, is 10 times larger, such that the average reproduction number of the total population remains constant. The average reproduction number can be calculated via

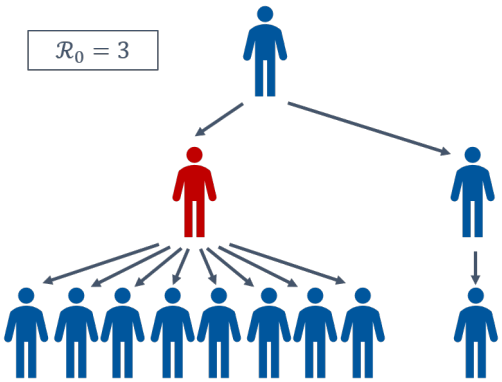


FIGURE 5.7: Visualization of overdispersion in secondary cases. Blue persons have a low level of contagiousness and transmit the disease to only a few other people, while the red person transmits the disease to many people. We refer to red persons as super spreaders. Anyone can become a super spreader.

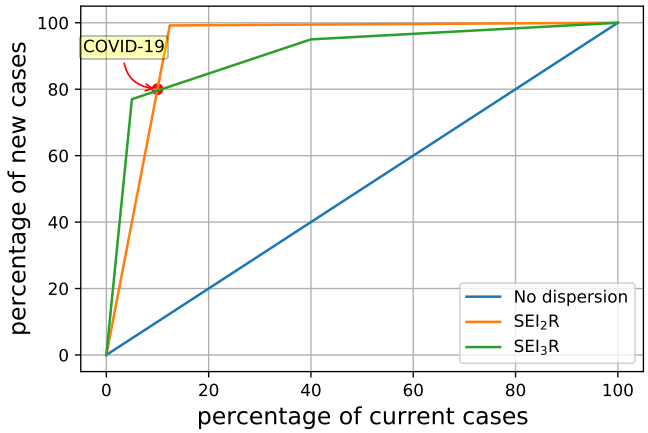


FIGURE 5.8: Percentage of new cases as a function of current cases. Our SEI_2R model follows the orange line. The SEI_3R model is added as an example.

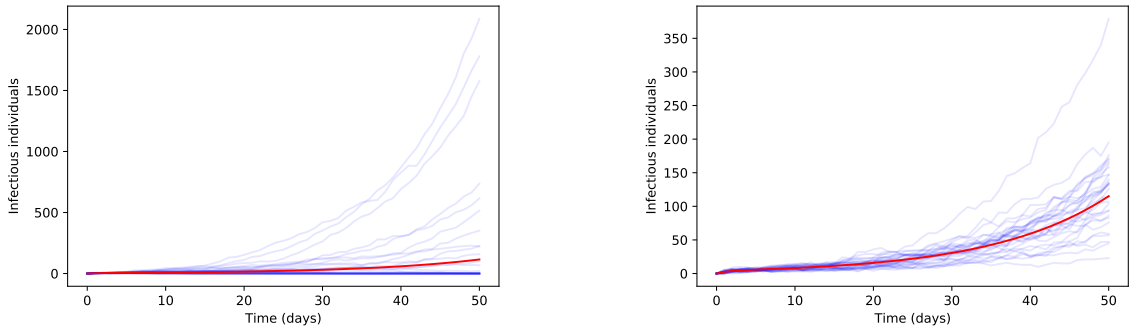
$$\mathcal{R}_{eff} = pR_{eff}^{\text{super}} + (1-p)R_{eff}^{\text{normal}}, \tag{5.10}$$

where $\mathcal{R}_{eff}^{\text{normal}}$ is the reproduction number of the other group, in our case, this is considered equal to 0. We use similar parameters as described in Table 5.1. This means that for $p = 0.1$, $\mathcal{R}_{eff}^{\text{super}} = 12.5$, such that, on average, the reproduction number equals 1.25, like in Table 5.1.

In Figure 5.9, we show a comparison between an SEI_2R model and an SEIR model when we initialize 10 individuals as exposed in a single municipality. All other individuals in The Netherlands are susceptible. The deterministic numerical result is colored red and equal for both models, as we expect. The blue lines represent individual stochastic simulations. We use stochastic rounding, as explained in Section 2.3.

By distinguishing the contagiousness levels of different individuals, we can capture overdispersion in secondary cases. Distinguishing between two levels of contagiousness already allows to approximate COVID-19's overdispersion. Distinguishing more levels allows to better model observed overdispersion [73].

We see that the effect of super spreaders has a large influence on the possible outcomes. The variability has increased. While we only use two different infected compartments, this could easily be extended to more classes; an SEI$_k$R model. Here, each infected compartment has its own level of contagiousness, and hence there own different reproduction numbers. This method could better match observations [29]. We leave further analysis to future work.



(a) SEI$_2$R model

(b) SEIR model

FIGURE 5.9: The effect of super spreaders (SEI$_2$R model) versus a normal SEIR model. Both models have equal $\mathcal{R}_0 = 1.25$, so the mean-field deterministic are equal. For SEI$_2$R model, the $\mathcal{R}_0^{\text{super}} = 12.5$ and $\mathcal{R}_0^{\text{normal}} = 0$ and the probability of becoming a super spreader equals $p = 0.1$. Initially, via our concentrated initialization, we initialize 10 individuals in the municipality of Uden as exposed, see Appendix F.

5.4 Evaluating mobility restrictions

For various initializations, we perform the optimization described in Section 5.2.5 and calculate the allowed mobility and the number of new infections for each of the strategies. Then we evaluate the results based on the objective function defined in Section 5.2.4. Throughout this section, we consider a time horizon of $T = 30$ days.

For each initialization, we apply the optimization method to various initial divisions. The initial divisions are the administrative divisions (provinces and security regions) and mobility-based initial divisions (mobility regions and adaptive mobility regions for various resolution parameters). Recall that we refrain from advising a specific value for the trade-off parameter since it is a political choice. However, in this section, we consider some values for α, but these are intended only to demonstrate the methodology.

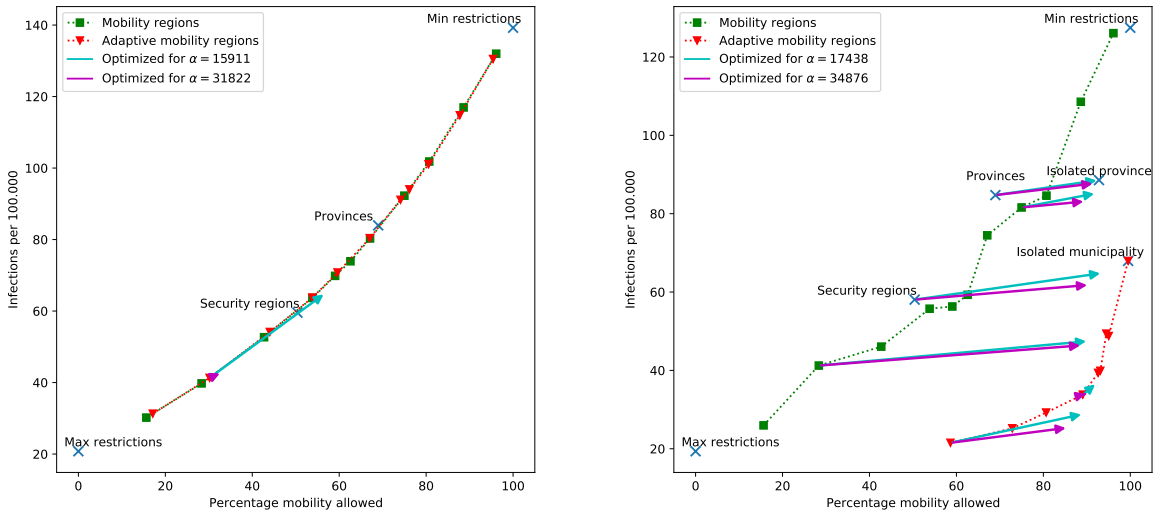
We consider two values for α: the value α^* for which maximum and minimum restrictions have equal objective value (i.e., $Q_{\alpha^*}(\mathcal{D}_{\min}) = Q_{\alpha^*}(\mathcal{D}_{\max})$ where $\mathcal{D}_{\max} = \mathcal{D}_{\min}$ are the divisions corresponding to maximum, respectively, minimum restrictions) and twice this value. Note that this value α^* is uniquely defined since $Q_{\alpha,T}(\mathcal{D})$ is a linear function in α.

We choose these values for the trade-off parameter rather than fixed constants, because constant trade-off parameters may lead to minimum restrictions for one initialization and maximum restrictions for another. This choice of α^* depends on the initialization and ensures a non-trivial division that outperforms both maximum and minimum restrictions. We start with the evenly distributed and concentrated initializations and draw a few observations. Then, we show to what extent these observations generalize to the historical initializations.

Results for evenly distributed and concentrated initializations. We consider an evenly distributed initialization where 1000 individuals are initially exposed, $e_x(t_0) \approx 6 \cdot 10^{-5}$ for all $x \in \mathcal{A}$. The resulting mobility and infections from each of the divisions are shown in Figure 5.10(a). We observe a monotonically increasing relation between infections and mobility. This may be explained by the fact that infections grow with increasing reproduction number, which is linear in the contact rate that in our model in turn linear in mobility. This trend is based on data points between 5000 and 30000 infections, which is less than an order of magnitude difference. The exponential-like trend is unlikely to hold for long time horizons. This will depend on the speed with which local saturation in contacts starts to influence transmission potential.

For the points shown in Figure 5.10(a), the division obtained by applying our optimization method to our mobility regions with the finest granularity (resolution parameter $= 16$, see Appendix G) resulted in the highest objective value for the trade-off parameter α^*. However, this division only performs marginally better than other divisions on this part of the curve. Furthermore, for a trade-off parameter of $2\alpha^*$, maximum restrictions performs best among the divisions shown.

Compared to provinces and security regions, we can use our mobility region's heuristics to tune the divisions' granularity. This allows us to find the optimal balance on this exponential-like curve with respect to the objective function for given values of the trade-off parameter and time horizon. We observe no divisions with significantly lower infection numbers in combination with as much mobility as in benchmark divisions. From this, we conclude that the amount of mobility that the division allows plays a larger role than the specific way this mobility is chosen.



(a) Evenly distributed

(b) Concentrated

FIGURE 5.10: Evaluation of different divisions for the evenly-distributed (left) versus concentrated (right) initialization. The balance of infections and mobility within a horizon of $T = 30$ days after imposing a division are shown. Blue crosses represent benchmark divisions. Green squares represent mobility regions, red triangles represent adaptive mobility regions; the resolution parameters are chosen to obtain divisions that allow various levels of mobility. Arrows point from the starting point of the heuristic optimization to its result. We only show a single arrow for the evenly distributed initialization to avoid cluttering the figure. The used values for α are shown in the legend.

If infections are not evenly distributed, the performances of the different divisions shift drastically. Figure 5.10(b) shows the performances of divisions for the concentrated initialization. In this initialization, 1000 people are initially exposed within the Dutch municipality of Uden (see Figure 5.6). In this case, $e_x(t_0) \approx 0.025$ only if $x =$ Uden. The mobility regions do not depend on the initialization of the model, so these divisions remain unchanged, see Appendix G. The adaptive mobility regions do depend on this initialization. We see that the adaptive mobility regions significantly outperform the mobility regions and the benchmark divisions: they result in fewer infections while allowing for more mobility.

If we apply our optimization method starting from a sufficiently fine-grained division, we obtain a division that performs even better. For both choices of the trade-off parameter α^* and $2\alpha^*$, the division obtained from applying the optimization to the adaptive mobility regions with the highest resolution (in other words, lowest percentage mobility allowed initially) outperforms all the other divisions shown with respect to the objective function.

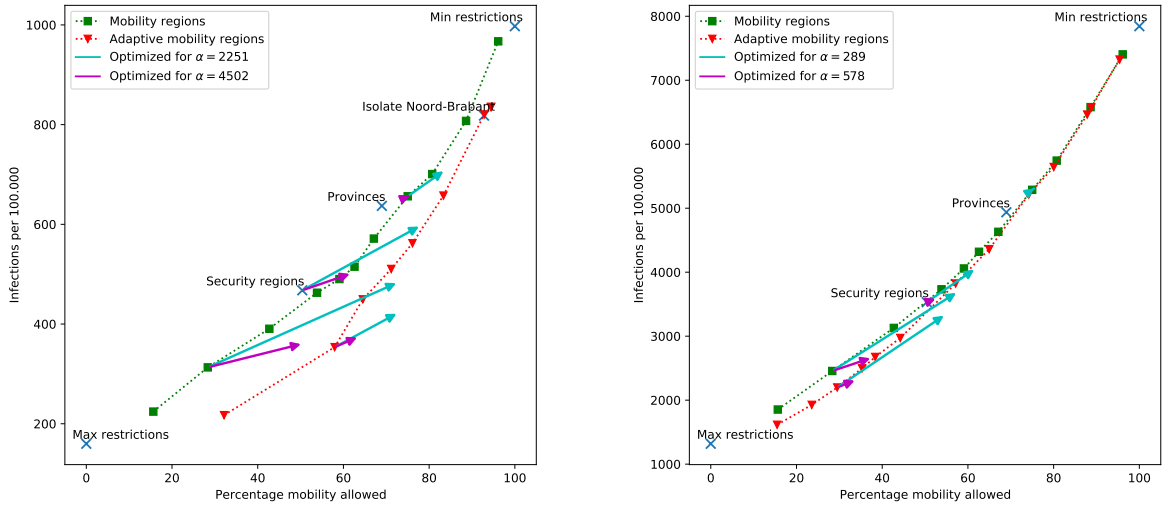
This suggests that the proposed approach can provide suitable strategies for containing super spreading events, where the initial infections are highly concentrated. This holds even when we change model parameters, see Appendix H, where we perform a sensitivity analysis.

Historical initializations. Next, we evaluate divisions for an initialization based on historical data. We consider data from March 10, 2020, and April 21, 2020. Figure 5.4 shows that the concentration values of historic initializations lie in between the concentration values of evenly distributed and concentrated cases. Therefore, we expect their results to also be in between the evenly distributed and concentrated results.

On March 10, 2020, the Dutch government advised all citizens of one of the twelve provinces to stay at home. At this point, the number of reported infections was low, and their distribution was far from even. April 21st was during the Dutch lock-down period, and further along in the first wave of the outbreak, where the infections were more evenly distributed.

Figures 5.11(a) and 5.11(b) show the evaluation results of different divisions. Figure 5.11(a) adds a local lock-down of the province containing Uden as a benchmark division. It performs almost as well as one of the divisions from the adaptive mobility regions approach. However, at that time, the Dutch government's goal was to suppress the virus in order to prevent new infections. From this perspective, only isolating this province is insufficient as it does not lead to significantly fewer infections than taking no action (Min restrictions).

Based on these two figures, we see that the findings from homogeneous and concentrated initializations generalize to more realistic scenarios. If the infections are concentrated in a small number of regions, our approach finds divisions that lead to relatively few infections, while allowing for relatively large percentages of mobility. Consider the marker corresponding to the security regions division in Figure 5.11(a): we can see that our mobility method has found divisions with 1) significantly more mobility, but a comparable number of infections; 2) a comparable percentage of allowed mobility, but significantly fewer infections; and 3) increased mobility and fewer infections. A similar conclusion can be made for the case of Figure 5.11(b), though the improvements are smaller because the initialization has a lower level of concentration.



(a) March 10 2020

(b) April 21 2020

FIGURE 5.11: Evaluation of divisions for initialization based on historical data of March 10 (left) versus April 21 (right) 2020. A horizon of $T = 30$ days was chosen. For our mobility regions and adaptive mobility regions, the resolution parameters are chosen to obtain divisions that allow various levels of mobility. The used values for α are shown in the legend.

5.5 Discussion

In a pandemic, restrictions on the mobility of individuals are one of the mitigation measures available to local and national governments. Restrictions on mobility between regions will have an effect on reducing non-local transmission opportunities. The downside, however, is that restricted mobility also has potentially strong social and economic repercussions. Given this, decision-makers have to reach a balance between wanted and unwanted effects if restricting mobility. Rather than impose restrictions on a national level, which could maximize unwanted effects, options need to be explored for regional measures. We have presented a method to determine a balance between infection reduction and allowed mobility. We evaluate mobility strategies that use any division of a country into regions, allowing movement within them while disallowing movement between them. In the case of The Netherlands, we have shown that existing (administrative) divisions such as provinces and 'security regions' do not reflect the mobility patterns within the country well, and therefore are not a good basis for mobility restrictions. We expect this conclusion to apply also to other countries.

We have quantified the trade-off between economy (equated here with mobility) and public health (equated here by infections) by introducing an objective function that penalizes the amount of allowed mobility by the resulting number of infections as given by Equation (5.8). This trade-off introduces a parameter that can be interpreted as the number of movements that we are willing to restrict in order to prevent the occurrence of a single further infection.

The objective function is non-linear, computationally heavy demanding, and hence infeasible to optimize exactly. Therefore, we resorted to heuristic optimization methods that are shown to produce divisions that perform well with respect to the presented objective.

As a proof of concept for the proposed methods, we have used scenarios with varying concentrations of infections. For each of these settings, we have compared how well the heuristics perform and have compared the divisions obtained in this way to the benchmarks of existing divisions. Figure 5.10(a) shows that when the infections are evenly distributed throughout the

country, the performances of all of the divisions lie close to an exponential curve. Therefore, the granularity of the division is the only relevant aspect in this case. On the other hand, Figure 5.10(b) shows that when the infections are highly concentrated around one municipality, applying the optimization to adaptive mobility regions results in a division that significantly outperforms the others and is able to prevent more infections while allowing for more mobility.

In practice, the spread of the infections will lie between these extremes. We have introduced a formula to quantify the concentration of the infections and Figures 5.11(a) and 5.11(b) show that indeed low concentration values lead to results comparable to the case of evenly distributed infections, while higher values indeed behave similarly as the concentrated case. Traditional epidemiological models do not incorporate geography and cannot adequately deal with situations where the infections are not evenly distributed across the country. In this work, we have demonstrated how incorporating geometry leads to mobility-restriction strategies that are better tailored to the situation at hand. Our main conclusion is that such strategies are highly effective when the geometric spread of infections is low (so that the geometry is the main limiting factor), but it is less effective when the distribution is rather even (so that geometry is fairly irrelevant). Our main innovation is that we have proposed a method to quantify these statements.

Chapter 6

Conclusion and recommendations

In this work, we have analyzed the dynamics of different spatial epidemic models to understand better what drives the spreading of epidemics. Motivated by the COVID-19 pandemic, we focused on the SIR and the SEIR compartmental models. We have demonstrated that epidemic models of the remote-force-of-infections class are well suited to model epidemics for mobility patterns of commuters. To our knowledge, analysis like ours of the SIR and the SEIR within the class of remote-force-of-infections models has not been done before.

Numerically, we have shown results on the spreading of epidemics for a one-dimensional network and two-dimensional networks. In all variations, we observe similar behavior after a localized outbreak of the disease. Depending on $\mathcal{R}_0$, there is diffusive growth which eventually transitions into a traveling wavefront of constant shape and velocity or diffusive decay, which transitions into decaying traveling waves. We expect that this behavior also holds for models with even more complex structures.

Analytically, we found asymptotic solutions that are valid shortly after a localized outbreak for a one-dimensional network. This results in a modified diffusion equation for the SIR model and a telegraph equation for the SEIR model. If the reproduction number is smaller than or equal to unity, $\mathcal{R}_0 \leq 1$, the telegraph equation describes the model behavior for larger values of t. We derived wave velocities for our traveling waves as a function model parameters, for $\mathcal{R}_0 > 1$. For the SIR model, the velocity is surprisingly similar to those found in reaction-diffusion epidemic models with similar parameters.

For COVID-19, the phenomena of super-spreaders play a role in the spreading of the disease. We proposed a method to include super-spreading effects into compartmental models, such as the SEIR model. We show that these effects, similar to other stochastic effects, only play a role when the number of individuals in the systems is small.

As an application of our SEIR model, we applied it to evaluate the impact of mobility restrictions in The Netherlands on the spread of COVID-19. For this, we investigated how to balance the trade-off between mobility restrictions and transmissions of COVID-19. We quantified the trade-off between blocking mobility and spreading the infection in terms of a single parameter. Our epidemic model evaluates mobility restrictions' effectiveness where the country is divided into separate regions. We propose heuristic methods to approximate optimal choices for these regions. The results show that our mobility restrictions are highly effective when the infections are concentrated around a few municipalities.

6.1 Future work

Understanding epidemic models help us understand the dynamics of epidemics. Asymptotic solutions help this understanding. If models become more accurate, they can also be applied to forecasting epidemics. We set out directions for future work.

Firstly, there is no work in the literature, to our knowledge, that directly compares compartmental models of the classes of reaction-diffusion epidemic models and remote-force-of-infection epidemic models. While both classes of models are still relevant, future work aimed at understanding differences between the two classes helps our understanding of modeling epidemics.

Secondly, future work could focus on an analytical description of the infectious profile $i(x,t)$ for the traveling wave regime. In our analytical analysis, we approximated the fraction susceptibilities by 1. This is valid for a short period after the initial outbreak and remains approximately valid if $\mathcal{R}_0 \leq 1$. For $\mathcal{R}_0 > 0$, we observe that the numerical solutions of our system transition into traveling waves, which are not captured by our asymptotic solutions.

Thirdly, it remains unclear if and for how long individuals remain resistant to COVID-19 after recovering from the disease. If resistance is lost eventually, the SEIRS model is a potential candidate for modeling COVID-19. We have left the analysis of the SEIRS and other compartmental models, such as SIS, SIRS, to future work. These models differ from the SIR and the SEIR models as they allow an infectious disease to become endemic. This means that there is a sustained fraction of individuals infectious, the epidemic does not die out.

Fourthly, we proposed a method that incorporated super-spreaders into compartmental models. The effect is only visible when the model is evaluated stochastically. Compartmental models are often used on applications at larger scales. For larger scales, stochastic effects become less relevant. Future work could aim to measure the impact of stochastic effects caused by super-spreaders. Models that include super-spreaders might be better at modeling or forecasting the spread of COVID-19.

Finally, in applying our SEIR model, we have decided to mostly base parameter choices relevant to the infection spread on previous studies. Alternatively, we could *estimate* these parameters by fitting the model to match historical data. By doing so, outcomes of the model may have a better predictive value. Currently, the number of infections is only used to compare with benchmark divisions. In particular, we cannot give estimates on how close these would be to the true value. Especially with historical scenarios, the fraction of reported cases and unreported cases depend heavily on the testing policy and availability of tests. Comparing RIVM estimates of active infections and reported infections hints that this fraction is indeed not at all constant [74]. In our model, we assume this fraction to be constant in all scenarios. It would be interesting to investigate the effect of heterogeneity, just like we now focus on the effect of heterogeneity of infections.

Bibliography

[1] Christophe Fraser et al. "Pandemic potential of a strain of influenza A (H1N1): early findings". In: *science* 324.5934 (2009), pp. 1557–1561.

[2] Vittoria Colizza et al. "Predictability and epidemic pathways in global outbreaks of infectious diseases: the SARS case study". In: *BMC medicine* 5.1 (2007), p. 34.

[3] Duygu Balcan et al. "Seasonal transmission potential and activity peaks of the new influenza A (H1N1): a Monte Carlo likelihood analysis based on human mobility". In: *BMC medicine* 7.1 (2009), p. 45.

[4] Duygu Balcan et al. "Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model". In: *Journal of computational science* 1.3 (2010), pp. 132–145.

[5] Steven Riley. "Large-scale spatial-transmission models of infectious disease". In: *Science* 316.5829 (2007), pp. 1298–1301.

[6] Roy M Anderson and Robert M May. *Infectious diseases of humans: dynamics and control*. Oxford university press, 1992.

[7] Yan Gang et al. "Epidemic spread in weighted scale-free networks". In: *Chinese Physics Letters* 22.2 (2005), p. 510.

[8] Dirk Brockmann and Dirk Helbing. "The hidden geometry of complex, network-driven contagion phenomena". In: *science* 342.6164 (2013), pp. 1337–1342.

[9] Martijn Gösgens et al. "Trade-offs between mobility restrictions and transmission of SARS-CoV-2". In: *Journal of the Royal Society Interface* 18.175 (2021), p. 20200936.

[10] William Ogilvy Kermack and Anderson G McKendrick. "A contribution to the mathematical theory of epidemics". In: *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character* 115.772 (1927), pp. 700–721.

[11] Odo Diekmann, Hans Heesterbeek, and Tom Britton. *Mathematical tools for understanding infectious disease dynamics*. Vol. 7. Princeton University Press, 2012.

[12] Phenyo E Lekone and Bärbel F Finkenstädt. "Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study". In: *Biometrics* 62.4 (2006), pp. 1170–1177.

[13] Ben M Bolker and Bryan Thomas Grenfell. "Chaos and biological complexity in measles dynamics". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 251.1330 (1993), pp. 75–81.

[14] Weston C Roda et al. "Why is it difficult to accurately predict the COVID-19 epidemic?" In: *Infectious Disease Modelling* (2020).

[15] Wildpixel. *Second Wave Virus Outbreak*. [Online; accessed February 18, 2021]. 2021. URL: https://www.gettyimages.nl/detail/foto/second-wave-virus-outbreak-royalty-free-beeld/1251262141.

[16] A Kolmogoroff, I Petrovsky, and N Piscounoff. "Study of the diffusion equation with growth of the quantity of matter and its application to a biology problem". In: *Dynamics of Curved Fronts*. Elsevier, 1988, pp. 105–130.

[17] Romualdo Pastor-Satorras et al. "Epidemic processes in complex networks". In: *Reviews of modern physics* 87.3 (2015), p. 925.

[18] Vittoria Colizza and Alessandro Vespignani. "Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations". In: *Journal of theoretical biology* 251.3 (2008), pp. 450–467.
[19] Alun L Lloyd and Robert M May. "Spatial heterogeneity in epidemic models". In: *Journal of theoretical biology* 179.1 (1996), pp. 1–11.
[20] Iikka A Hanski, Michael E Gilpin, and David E McCauley. *Metapopulation biology*. Vol. 454. Elsevier, 1997.
[21] Neil M Ferguson et al. "Strategies for containing an emerging influenza pandemic in Southeast Asia". In: *Nature* 437.7056 (2005), pp. 209–214.
[22] Neil M Ferguson et al. "Strategies for mitigating an influenza pandemic". In: *Nature* 442.7101 (2006), pp. 448–452.
[23] Stefano Merler and Marco Ajelli. "The role of population heterogeneity and human mobility in the spread of pandemic influenza". In: *Proceedings of the Royal Society B: Biological Sciences* 277.1681 (2010), pp. 557–565.
[24] Philip M Morse and Herman Feshbach. "Methods of theoretical physics". In: *American Journal of Physics* 22.6 (1954), pp. 410–413.
[25] VH Weston and S He. "Wave splitting of the telegraph equation in R3 and its application to inverse scattering". In: *Inverse Problems* 9.6 (1993), p. 789.
[26] PM Jordan and Ashok Puri. "Digital signal propagation in dispersive media". In: *Journal of Applied Physics* 85.3 (1999), pp. 1273–1282.
[27] Willy Malfliet. "Solitary wave solutions of nonlinear wave equations". In: *American journal of physics* 60.7 (1992), pp. 650–654.
[28] Lei Dong et al. "Measuring economic activity in China with mobile big data". In: *EPJ Data Science* 6.1 (2017), p. 29.
[29] Agus Hasan et al. "Superspreading in early transmissions of COVID-19 in Indonesia". In: *Scientific reports* 10.1 (2020), pp. 1–4.
[30] Alison L. Hill. "The math behind epidemics". In: *Physics Today* 73.28 (2020).
[31] Roy M Anderson et al. "How will country-based mitigation measures influence the course of the COVID-19 epidemic?" In: *The Lancet* 395.10228 (2020), pp. 931–934.
[32] William O Kermack and Anderson G McKendrick. "Contributions to the mathematical theory of epidemics–I. 1927." In: *Bulletin of mathematical biology* 53.1-2 (1991), pp. 33–55.
[33] George Gee Jackson et al. "Susceptibility and immunity to common upper respiratory viral infections—the common cold". In: *Annals of Internal Medicine* 53.4 (1960), pp. 719–738.
[34] Brian J Coburn, Bradley G Wagner, and Sally Blower. "Modeling influenza epidemics and pandemics: insights into the future of swine flu (H1N1)". In: *BMC medicine* 7.1 (2009), pp. 1–8.
[35] Mevin B Hooten, Jessica Anderson, and Lance A Waller. "Assessing North American influenza dynamics with a statistical SIRS model". In: *Spatial and spatio-temporal epidemiology* 1.2-3 (2010), pp. 177–185.
[36] Steven M Goodreau et al. "Concurrent partnerships, acute infection and HIV epidemic dynamics among young adults in Zimbabwe". In: *AIDS and Behavior* 16.2 (2012), pp. 312–322.
[37] Mariusz Bodzioch, Marcin Choiński, and Urszula Foryś. "SIS criss-cross model of tuberculosis in heterogeneous population". In: *Discrete & Continuous Dynamical Systems-B* 24.5 (2019), p. 2169.
[38] JD Hernández Guillén, A Martın Del Rey, and L Hernández Encinas. "Study of the stability of a SEIRS model for computer worm propagation". In: *Physica A: Statistical Mechanics and its Applications* 479 (2017), pp. 411–421.

[39] Robert D Kirkcaldy, Brian A King, and John T Brooks. "COVID-19 and postinfection immunity: limited evidence, many remaining questions". In: *Jama* 323.22 (2020), pp. 2245–2246.
[40] Haley E Randolph and Luis B Barreiro. "Herd immunity: understanding COVID-19". In: *Immunity* 52.5 (2020), pp. 737–741.
[41] World Health Organization et al. *Transmission of SARS-CoV-2: implications for infection prevention precautions: scientific brief, 09 July 2020*. Tech. rep. World Health Organization, 2020.
[42] Tiberiu Harko, Francisco SN Lobo, and MK Mak. "Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates". In: *Applied Mathematics and Computation* 236 (2014), pp. 184–194.
[43] Nicola Piovella. "Analytical solution of SEIR model describing the free spread of the COVID-19 pandemic". In: *Chaos, Solitons & Fractals* 140 (2020), p. 110243.
[44] *Effective Reproduction Number as provided by the Dutch Government*. https://coronadashboard.rijksoverheid.nl/landelijk/reproductiegetal. Accessed: 2021-03-21.
[45] Odo Diekmann, Johan Andre Peter Heesterbeek, and Johan AJ Metz. "On the definition and the computation of the basic reproduction ratio R 0 in models for infectious diseases in heterogeneous populations". In: *Journal of mathematical biology* 28.4 (1990), pp. 365–382.
[46] Qun Li et al. "Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia". In: *New England Journal of Medicine* (2020).
[47] Daniel Simberloff. "The contribution of population and community biology to conservation science". In: *Annual review of ecology and systematics* 19.1 (1988), pp. 473–511.
[48] Vittoria Colizza et al. "The role of the airline transportation network in the prediction and predictability of global epidemics". In: *Proceedings of the National Academy of Sciences* 103.7 (2006), pp. 2015–2020.
[49] Romualdo Pastor-Satorras and Alessandro Vespignani. "Epidemic dynamics and endemic states in complex networks". In: *Physical Review E* 63.6 (2001), p. 066117.
[50] Filippo Simini et al. "A universal model for mobility and migration patterns". In: *Nature* 484.7392 (2012), pp. 96–100.
[51] VV Belik and Dirk Brockmann. "Accelerating random walks by disorder". In: *New Journal of Physics* 9.3 (2007), p. 54.
[52] S Rushton and AJ Mautner. "The deterministic model of a simple epidemic for more than one community". In: *Biometrika* 42.1/2 (1955), pp. 126–132.
[53] Matt J Keeling and Pejman Rohani. "Estimating spatial coupling in epidemiological systems: a mechanistic approach". In: *Ecology Letters* 5.1 (2002), pp. 20–29.
[54] Duygu Balcan et al. "Multiscale mobility networks and the spatial spreading of infectious diseases". In: *Proceedings of the National Academy of Sciences* 106.51 (2009), pp. 21484–21489.
[55] Odo Diekmann, JAP Heesterbeek, and Michael G Roberts. "The construction of next-generation matrices for compartmental epidemic models". In: *Journal of the Royal Society Interface* 7.47 (2010), pp. 873–885.
[56] Richard V Kadison and John R Ringrose. *Fundamentals of the theory of operator algebras*. Vol. 1. American Mathematical Soc., 1983, pp. 204–205.
[57] Andrew W Roddam. *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation: O Diekmann and JAP Heesterbeek, 2000, Chichester: John Wiley pp. 303,£ 39.95. ISBN 0-471-49241-8*. 2001.
[58] Odo Diekmann, JAP Heesterbeek, and Tom Britton. *Mathematical Tools for Understanding Infectious Disease Dynamics:* Princeton University Press, 2013. ISBN: 9780691155395. URL: http://www.jstor.org/stable/j.cttq9530.
[59] Rick Durrett. *Probability: theory and examples*. Vol. 49. Cambridge university press, 2019.
[60] Philip G Drazin and Robin S Johnson. *Solitons: an introduction*. Vol. 2. Cambridge university press, 1989, p. 15.

[61] Henry C Tuckwell. "Evidence of soliton-like behavior of solitary waves in a nonlinear reaction-diffusion system". In: *SIAM Journal on Applied Mathematics* 39.2 (1980), pp. 310–322.
[62] Gerald B Folland. *Introduction to partial differential equations*. Vol. 102. Princeton university press, 1995.
[63] Edward A Codling, Michael J Plank, and Simon Benhamou. "Random walk models in biology". In: *Journal of the Royal society interface* 5.25 (2008), pp. 813–834.
[64] Gerald B Folland. *Fourier analysis and its applications*. Vol. 4. American Mathematical Soc., 2009.
[65] Ronald Aylmer Fisher. "The wave of advance of advantageous genes". In: *Annals of eugenics* 7.4 (1937), pp. 355–369.
[66] Vitaly Belik, Theo Geisel, and Dirk Brockmann. "Natural human mobility patterns and spatial spread of infectious diseases". In: *Physical Review X* 1.1 (2011), p. 011001.
[67] Wim van Saarloos. "Front propagation into unstable states". In: *Physics Reports* 386.2 (2003), pp. 29–222. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2003.08.001. URL: https://www.sciencedirect.com/science/article/pii/S0370157303003223.
[68] Robin N Thompson et al. "Key questions for modelling COVID-19 exit strategies". In: *Proceedings of the Royal Society B* 287.1932 (2020), p. 20201405.
[69] Joël Mossong et al. "Social contacts and mixing patterns relevant to the spread of infectious diseases". In: *PLoS Med* 5.3 (2008), e74.
[70] *COVID-19 cumulative numbers of positive tested cases per municipality in The Netherlands*. https://data.rivm.nl/geonetwork/srv/dut/catalog.search#/metadata/1c0fcd57-1102-4620-9cfa-441e93ea5604?tab=relations. Accessed: 2021-03-21.
[71] Yuhao Deng et al. "Estimation of incubation period and generation time based on observed length-biased epidemic cohort with censoring for COVID-19 outbreak in China". In: *Biometrics* (2020).
[72] Natalie M Linton et al. "Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: a statistical analysis of publicly available case data". In: *Journal of clinical medicine* 9.2 (2020), p. 538.
[73] Akira Endo et al. "Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China". In: *Wellcome Open Research* 5 (2020).
[74] Rijksinstituut voor Volksgezondheid en Milieu - RIVM. *Wekelijkse update epidemiologische situatie COVID-19 in Nederland*. 2020. URL: https://www.rivm.nl/en/media/158941.
[75] Jeffrey I Steinfeld, Joseph Salvadore Francisco, and William L Hase. *Chemical kinetics and dynamics*. Prentice Hall Upper Saddle River, NJ, 1999.
[76] Peter Schuster. "What is special about autocatalysis?" In: *Monatshefte für Chemie-Chemical Monthly* 150.5 (2019), pp. 763–775.
[77] Vincent A Traag, Ludo Waltman, and Nees Jan van Eck. "From Louvain to Leiden: guaranteeing well-connected communities". In: *Scientific reports* 9.1 (2019), pp. 1–12.
[78] Santo Fortunato. "Community detection in graphs". In: *Physics reports* 486.3-5 (2010), pp. 75–174.
[79] Theus Hossmann, Thrasyvoulos Spyropoulos, and Franck Legendre. "A complex network analysis of human mobility". In: *2011 IEEE conference on computer communications workshops (INFOCOM WKSHPS)*. IEEE. 2011, pp. 876–881.
[80] Mehmet Yildirimoglu and Jiwon Kim. "Identification of communities in urban mobility networks using multi-layer graphs of network traffic". In: *Transportation Research Part C: Emerging Technologies* 89 (2018), pp. 254–267.
[81] Santo Fortunato and Marc Barthelemy. "Resolution limit in community detection". In: *Proceedings of the national academy of sciences* 104.1 (2007), pp. 36–41.
[82] Vincent D Blondel et al. "Fast unfolding of communities in large networks". In: *Journal of statistical mechanics: theory and experiment* 2008.10 (2008), P10008.

Appendix A

Model variations

In this chapter, we describe epidemic models from Table 2.1 that are not discussed in the main text. We recommend reading Section 2.1.2 before reading this appendix. The models are briefly discussed in Section 2.1.4. Here, we present the systems of equations that describe the SI, the SIS, the SIRS, and the SEIRS model for a single population.

SI model. The Susceptible-Infectious model has two compartments, susceptible and infectious. Once infected, you are immediately contagious and remain contagious.

$$\frac{ds}{dt} = -\beta si, \tag{A.1}$$

$$\frac{di}{dt} = \beta si = -\frac{ds}{dt}. \tag{A.2}$$

From these equations, we can see similarities between this epidemic model and, for example, a model for autocatalysis reactions [75]. In a simple case, we have two kinds of particles in solution: particles of type A and B. They react via $A + B \rightarrow 2B$ with a reaction constant k. We can write the ODE governing this reaction as:

$$\frac{d[A]}{dt} = -k[A][B] \tag{A.3}$$

where $[\cdot]$ indicates the concentration of a certain particle type. (A.3) is very similar to equation A.1. From this field, we can use the exact solution for $s(t)$, which equals [76]

$$i(t) = \frac{i(t_0)}{i(t_0) + s(t_0)\exp\left[-\beta t\right]}, \tag{A.4}$$

where $s(t_0)$ and $i(t_0)$ are initial conditions and the conservation law equals $1 = s + i$

SIS model. The Susceptible-Infectious-Susceptible model has two compartments, susceptible and infectious. Once infected, you are immediately contagious. The infectious period has average an duration $1/\gamma$. Instead of recovering after this period, individuals become susceptible again. We can write the ODEs for this SIS model as follows:

$$\frac{ds}{dt} = -\beta si + \gamma i, \tag{A.5}$$
$$\frac{di}{dt} = \beta si - \gamma i = -\frac{ds}{dt}. \tag{A.6}$$

If we make time dimensionless, as we did in section 2.1.2, we find $\mathcal{R}_0 = \beta/\gamma$, similar to the SIR model. To simplify our system of ODEs, we can use $I = N - S$, or in densities, $i = 1 - s$ to reduces the system to a single equation:

$$\frac{di}{dt} = (\beta - \gamma)i + \beta i^2. \tag{A.7}$$

From here we can easily calculate the two equilibrium points, at $i(t) = 0$ and if $\mathcal{R}_0 > 1$ we also have $i(t) = \frac{\beta-\gamma}{\beta}$.

SIRS model The Susceptible-Infectious-Recovered-Susceptible model has three compartments, susceptible, infected and recovered. Once infected, individuals are immediately contagious. After recovery, individuals lose their resistance to the disease with rate α and become susceptible again. We can write the ODEs for this SIRS model as follows:

$$\frac{ds}{dt} = -\beta si + \alpha r, \tag{A.8}$$
$$\frac{di}{dt} = \beta si - \gamma i, \tag{A.9}$$
$$\frac{di}{dt} = \gamma i - \alpha r. \tag{A.10}$$

If we make time dimensionless, as we did in section 2.1.2, we find $\mathcal{R}_0 = \beta/\gamma$, similar to the SIR model. The conservation law equals $1 = s + i + r$ and the system has initial conditions $s(t_0)$, $i(t_0)$, and $r(t_0)$.

SEIRS model The Susceptible-Exposed-Infectious-Recovered-Susceptible model has four compartments, susceptible, exposed, infected, and recovered. Once infected, individuals become exposed to the disease. With rate ν, individuals become contagious. The average time spent in the exposed compartment is called the latent period, $1/\nu$. After recovery, individuals lose their resistance to the disease with rate α and become susceptible again. The average period in which individuals are resistant to the disease is of length $1/\alpha$. We can write the ODEs for this SIRS model as follows:

$$\frac{ds}{dt} = -\beta si + \alpha r, \tag{A.11}$$
$$\frac{de}{dt} = \beta si - \nu e, \tag{A.12}$$
$$\frac{di}{dt} = \nu e - \gamma i, \tag{A.13}$$
$$\frac{dr}{dt} = \gamma i - \alpha r. \tag{A.14}$$

If we make time dimensionless, as we did in section 2.1.2, we find $\mathcal{R}_0 = \beta/\gamma$, similar to the SIR model. The conservation law equals $1 = s + e + i + r$ and the system has initial conditions $s(t_0)$, $e(t_0)$, $i(t_0)$, and $r(t_0)$.

Appendix B

Epidemic final-size derivation for SIR model

Here, we derive the final-size Equation (2.22) as presented in Section 2.1.4. We use Equations 2.7 and 2.9. In Section 2.1.4, we argued that for $t \to \infty$, the infectious concentration goes to $i(t) \to 0$.

Firstly, we divide Equation (2.7) by Equation (2.9). We find

$$\frac{ds}{dr} = -\mathcal{R}_0 s. \tag{B.1}$$

We rewrite this into

$$\frac{1}{s} ds = -\mathcal{R}_0 dr. \tag{B.2}$$

We integrate both sides from $t = 0$ to $t = \infty$ to find

$$\ln\left[\frac{s_\infty}{s(t=0)}\right] = -\mathcal{R}_0(r_\infty - r(t=0)). \tag{B.3}$$

When we initialize the system as $s(t=0) = 1 - i_0$, $i(t=0) = i_0$, and $r(t=0) = 0$. For $t \to \infty$, we write $s(t=\infty) = s_\infty$ and $r(t=\infty) = r_\infty$. Using that $i_\infty = 0$ and the conservation law, we rewrite $s_\infty = 1 - r_\infty$. Then, we find

$$\ln\left[\frac{1 - r_\infty}{1 - i_0}\right] = -\mathcal{R}_0(r_\infty). \tag{B.4}$$

We rewrite this as

$$r_\infty = 1 - (1 - i_0)\exp\left[-\mathcal{R}_0(r_\infty)\right]. \tag{B.5}$$

which is equal to Equation (2.22).

Appendix C

Modified diffusion equation

Here, we solve (3.27) using Green's functions [24]. We replace $i(x,t)$ with $u(x,t)$ to prevent confusion with the imaginary number i, which we encounter in finding the solution.

$$\frac{\partial u(x,t)}{\partial t} = \frac{\partial^2 u(x,t)}{\partial x^2} + (1-b)u(x,t). \tag{C.1}$$

We use the the wave propagation convention for the Fourier transform [64]:

$$u(k,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u(x,t)\exp\left[-ikx\right]dx, \tag{C.2}$$

and its corresponding inverse Fourier transform:

$$u(x,t) = \int_{-\infty}^{\infty} u(k,t)\exp\left[ikx\right]dx. \tag{C.3}$$

We substitute the inverse Fourier transform of $u(x,t)$ in (C.1). We find:

$$\frac{\partial u(k,t)}{\partial t} = (1-b-k^2)u(k,t). \tag{C.4}$$

We know that solution of this differential equation is of the shape:

$$u(k,t) = u(k,t=t')\exp\left[(1-b-k^2)(t-t')\right]. \tag{C.5}$$

Now, we assume that we have a point source at $t = t'$ located at $x = x'$. For the point source, we use a Dirac delta distribution $\delta(x)$. We find $u(x,t=t') = \delta(x-x')$. The Fourier transform of $u(x,t=t')$ is equal to the Fourier transform of a Dirac delta distribution:

$$u(k,t=t') = \frac{1}{2\pi}\int_{-\infty}^{\infty}\delta(x-x')\exp\left[ikx\right]dx = \frac{\exp\left[-ikx'\right]}{2\pi}. \tag{C.6}$$

We find u(x,t) by substituting calculating the Fourier transform of $u(k,t)$ in Equation (C.5). We substitute the result for $u(k,t=t')$ in Equation (C.5). We calculate the Fourier transform. We find

$$u(x,x',t,t') = \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\left[ik(x-x')\right]\exp\left[(1-b-k^2)(t-t')\right]dk. \tag{C.7}$$

The solution to this equation is equal to Green's function [24]. We can solve this integral by *completing the squares*:

$$G(x,x',t,t') = \frac{1}{2\pi}\exp\left[(1-b)(t-t')\right]\int_{-\infty}^{\infty}\exp\left[ik(x-x')\right]\exp\left[-k^2(t-t')\right]dk, \tag{C.8}$$

$$G(x,x',t,t') = \frac{1}{2\pi}\exp\left[(1-b)(t-t')\right]\exp\left[-\frac{(x-x')^2}{4(t-t')}\right]\int_{-\infty}^{\infty}\exp\left[-(t-t')\left(k-\frac{x-x'}{2(t-t')}\right)^2\right]dk, \tag{C.9}$$

$$G(x,x',t,t') = \frac{1}{2\pi}\exp\left[(1-b)(t-t')\right]\exp\left[-\frac{(x-x')^2}{4(t-t')}\right]\int_{-\infty}^{\infty}\exp\left[-(t-t')\,(z)^2\right]dz, \tag{C.10}$$

$$G(x,x',t,t') = \frac{1}{2\pi}\exp\left[(1-b)(t-t')\right]\exp\left[-\frac{(x-x')^2}{4(t-t')}\right]\sqrt{\frac{\pi}{t-t'}}, \tag{C.11}$$

$$G(x,x',t,t') = \frac{1}{\sqrt{4\pi(t-t')}}\exp\left[-\frac{(x-x')^2}{4(t-t')}\right]\exp\left[(1-b)(t-t')\right]. \tag{C.12}$$

Now, we can calculate $u(x,t)$ if we use the initial condition $u(x,0)$ and Green's function via:

$$u(x,t) = \int_{-\infty}^{\infty} G(x,x',t,t')u(x',t=0)dx'. \tag{C.13}$$

For $u(x,t=0) = \delta(x-x_0)$, we find the solution to (C.1):

$$u(x,t) = \frac{1}{\sqrt{4\pi t}}\exp\left[-\frac{(x-x_0)^2}{4t}\right]\exp\left[(1-b)t\right]. \tag{C.14}$$

Appendix D

Telegraph equation solution

Here, we solve (3.30). We replace $i(x,t)$ with $u(x,t)$ to prevent confusion with the imaginary number i, which we encounter in finding the solution. The PDE from Equation (3.30) reads

$$\frac{\partial^2 u(x,t)}{\partial t^2} + \left(\frac{1}{\mathcal{R}_0} + c\right)\frac{\partial u(x,t)}{\partial t} - c\frac{\partial^2 u(x,t)}{\partial x^2} = c\left(1 - \frac{1}{\mathcal{R}_0}\right)u(x,t). \tag{D.1}$$

We consider the following initial and boundary conditions for the domain $x = [0, L]$ and $t = [0, \infty)$:

$$i(x,0) = i_0\delta(x - x_0), \tag{D.2}$$
$$\frac{\partial i}{\partial t} = 0, \tag{D.3}$$
$$i(0,t) = 0, \tag{D.4}$$
$$i(L,t) = 0. \tag{D.5}$$

We find the solution using the method of separation of variables [64]. This means that we write $u(x,t) = X(x)T(t)$. We define the parameter $b = 1/\mathcal{R}_0$. By substituting this in Equation (D.1), we find

$$XT'' + (b + c)\,xT' + c(b-1)XT = cTX''. \tag{D.6}$$

Separating the variables, we find

$$\frac{X''}{X} = \frac{1}{c}\left(\frac{T''}{T} + (b+c)\frac{T'}{T} + (b-1)\right) = \sigma \tag{D.7}$$

We use boundary conditions (D.4) and (D.5) to find an expressions for $X(x)$:

$$X(x=0) = 0, \qquad X(x=L) = 0, \qquad X(x) = c_1 \sin\left(\frac{n\pi x}{L}\right), \qquad \sigma = -\left(\frac{n\pi}{L}\right)^2$$

The ODE for $T(t)$ equals

$$T'' + (b+c)T' + (b - 1 - c\sigma)T = 0 \tag{D.8}$$

This ODE has a general solution of the form

$$T(t) = c_2 \exp\left[\lambda_1 t\right] + c_3 \exp\left[\lambda_2 t\right]. \tag{D.9}$$

In this general solution, λ_1 and λ_2 are roots of the equation

$$\lambda^2 + (b+c)\lambda + (b - 1 - c\sigma) = 0, \tag{D.10}$$

which are equal to

$$\lambda_{1,2} = \frac{1}{2}(-(b+c) \pm \sqrt{(b+c)^2 - 4(b-1-c\sigma)}). \tag{D.11}$$

Let $4\omega_n^2$ be equal to $4(b-1-c\sigma) - (b+c)^2$. Then we write

$$\lambda_{1,2} = -\frac{b+c}{2} \pm i\omega_n. \tag{D.12}$$

We substitute this into Equation (D.9), so that we can write $T(t)$ as

$$T(t) = A_n \exp\left[-\frac{b+c}{2}t\right] \cos\left(\omega_n t - \phi_n\right), \tag{D.13}$$

which has an amplitude equal to A_n and an arbitrary phase ϕ_n. This means that for all amplitudes and phases, we have

$$u(x,t) = \sum_{n=1}^{\infty} A_n \exp\left[-\frac{b+c}{2}t\right] \cos\left(\omega_n t - \phi_n\right) \sin\left(\frac{n\pi x}{L}\right). \tag{D.14}$$

Equation (D.14) satisfies the Equation (D.1) and the two boundary conditions (D.4) and (D.5). The initial conditions (D.2) and (D.3) determine the amplitude and the phase for different values of n.

For boundary condition (D.3), we need to use the derivative of $u(x,t)$,

$$\frac{\partial u(x,t)}{\partial t} = \sum_{n=1}^{\infty} A_n \exp\left[-\frac{b+c}{2}t\right] \left(-\frac{b+c}{2}\cos\left(\omega_n t - \phi_n\right) - \omega_n \sin\left(\omega_n t - \phi_n\right)\right) \sin\left(\frac{n\pi x}{L}\right). \tag{D.15}$$

Then, we use the boundary condition to find

$$\frac{\partial u(x,0)}{\partial t} = \sum_{n=1}^{\infty} A_n \left(-\frac{b+c}{2}\cos\left(\phi_n\right) + \omega_n \sin\left(\phi_n\right)\right) \sin\left(\frac{n\pi x}{L}\right) = 0, \tag{D.16}$$

$$= -\frac{b+c}{2}\cos\left(\phi_n\right) + \omega_n \sin\left(\phi_n\right) = 0 \tag{D.17}$$

The latter is the case when we chose $\phi_n = \arctan\left(\frac{b+c}{2\omega_n}\right)$. Lastly, we use boundary (D.2) to find

$$u(x,0) = \sum_{n=1}^{\infty} A_n \cos\left(\phi_n\right) \sin\left(\frac{n\pi x}{L}\right) = \delta(x - x_0), \tag{D.18}$$

$$A_n \cos\left(\phi_n\right) = \frac{2}{L}\int_0^L \delta(x-x_0) \sin\left(\frac{n\pi x}{L}\right) = \sin\left(\frac{n\pi x_0}{L}\right). \tag{D.19}$$

This results in

$$i(x,t) = \frac{1}{2}\sum_{n=1}^{\infty} A_n \exp\left[-\left(\frac{c}{2} + \frac{1}{2\mathcal{R}_0}\right)t\right] \left(\sin\left(\frac{n\pi x}{L} - \omega_n t + \phi_n\right) + \sin\left(\frac{n\pi x}{L} + \omega_n t - \phi_n\right)\right), \tag{D.20}$$

where the constants are as follows:

$$\omega_n = \sqrt{c\left(\left(\frac{n\pi}{L}\right)^2 - \frac{5}{4} - \frac{1}{4c\mathcal{R}_0}\right)}, \qquad \phi_n = \arctan\left(\frac{c}{2\omega_n} + \frac{1}{2\omega_n \mathcal{R}_0}\right), \qquad A_n = \frac{2}{L\cos\phi_n}\sin\left(\frac{n\pi x_0}{L}\right).$$

Appendix E

Reaction diffusion model

Here we briefly show the differences between remote-force-of-infection models and reaction diffusion epidemic models for a metapopulation network similar to Figure 3.1. For the SIR remote-force-of-infection model, the one-dimensional equations are presented in Section 3.1. We also put them here for convenience:

$$\frac{\partial s(x,t)}{\partial t} = -s(x,t)i(x,t) - s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2}, \tag{E.1}$$

$$\frac{\partial i(x,t)}{\partial t} = s(x,t)i(x,t) + s(x,t)\frac{\partial^2 i(x,t)}{\partial x^2} - \frac{1}{\mathcal{R}_0}i(x,t), \tag{E.2}$$

$$\frac{\partial r(x,t)}{\partial t} = \frac{1}{\mathcal{R}_0}i(x,t). \tag{E.3}$$

Here, the basic reproduction number $\mathcal{R}_0$ is the only remaining parameter and $\mathcal{R}_0 = \frac{\beta_l + 4m\beta_m}{\gamma}$. Together with initial conditions ,$s(x, t = 0)$, $i(x, t = 0)$, and $r(x, t = 0)$, and the conservation law $r(x,t) = 1 - s(x,t) - i(x,t)$, the system of PDEs in eq. (3.4) and eq. (3.5) forms a well-defined system.

Now, for Equations 2.25, 2.26, and 2.27 in combination with a network similar to Figure 3.1, we obtain different equations. We assume that the mobility rates ω_{xy} are all equal $\omega_{xy} \equiv \omega$, for simplicity. We transform time as $t = \frac{1}{\beta}t^*$ and the spatial coordinate in our network as $x = \sqrt{\frac{\omega\Delta x^2}{\beta}}x^*$. We can write down the dimensionless PDEs for an SIR reaction diffusion model as follows:

$$\frac{\partial s(x,t)}{\partial t} = -s(x,t)i(x,t) + \frac{\partial^2 s(x,t)}{\partial x^2}, \tag{E.4}$$

$$\frac{\partial i(x,t)}{\partial t} = s(x,t)i(x,t) - \frac{1}{\mathcal{R}_0}i(x,t) + \frac{\partial^2 i(x,t)}{\partial x^2}, \tag{E.5}$$

$$\frac{\partial r(x,t)}{\partial t} = \frac{1}{\mathcal{R}_0}i(x,t) + \frac{\partial^2 r(x,t)}{\partial x^2}. \tag{E.6}$$

Here, we have dropped the asterisks, and the basic reproduction number $\mathcal{R}_0$ is the only remaining parameter, and $\mathcal{R}_0 = \frac{\beta}{\gamma}$. The conservation law equals $1 = s(x,t) + i(x,t) + r(x,t)$. Together with initial conditions, the system of PDEs forms a well-defined system.

Appendix F

Model initialization and the concentration measure

Here, we present details about the numerical work from Chapter 5. This includes model initialization and the calculation of the concentration measure.

F.1 Model initializations

In this section, we discuss the details of the different initializations as discussed in Chapter 5.

Evenly-distributed initialization. A fixed fraction e_0 of people in all municipalities is set to exposed, i.e.,

$$s_x(t_0) = (1 - e_0), \qquad e_x(t_0) = e_0, \qquad i_x(t_0) = 0, \qquad r_x(t_0) = 0.$$

Expressed in number of individuals, we have

$$S_x(t_0) = (1 - e_0) N_x, \qquad E_x(t_0) = e_0 N_x, \qquad I_x(t_0) = 0, \qquad R_x(t_0) = 0.$$

Note that the number of people who are susceptible does not need to be integer-valued in a deterministic simulation. For stochastic simulation, we round all values to the closest integer. After initializing, we simulate for 10 days to mimic that the virus has spread to other municipalities before it is found and people are tested. The resulting state at $t = t_0 + 10$ can then be used as an input for the adaptive mobility regions approach (see Appendix G) and also as a starting point for evaluating the performance of different divisions.

Concentrated initialization. The idea behind a concentrated initialization is to create the situations caused by a superspreader. Where locally, in a single municipality, there are many individuals exposed, but individuals from other all municipalities are still all susceptible to the disease. The fraction that we initially expose in a single municipality equals e_0. We initialize this in municipality x_0. We use the Kronecker delta $\delta_{x,y}$ notation, which is only equal to one when $x = y$ and zero else:

$$s_x(t_0) = (1 - e_0\delta_{x,x0}), \qquad e_x(t_0) = e_0\delta_{x,x0}, \qquad i_x(t_0) = 0, \qquad r_x(t_0) = 0.$$

Expressed in number of individuals, we have

$$S_x(t_0) = (1 - e_0\delta_{x,x0}) N_x, \qquad E_x(t_0) = e_0\delta_{x,x0} N_x, \qquad I_x(t_0) = 0, \qquad R_x(t_0) = 0.$$

Similarly, in the case of evenly-distributed initializations, we initialize by rounding all values to the closest integer. After initializing, we simulate for 10 days to mimic that the virus has spread to other municipalities before it is found and people are tested. The resulting state at $t = t_0 + 10$ is then used as input for our approach.

Initialization based on historic data. For an initialization based on historical situations, RIVM data is used (RIVM is the Dutch institute of public health). RIVM reports the cumulative number of infections for each municipality. The number of currently active known infections at time t_0, is equal to the reported infections at time t_0, $I_i^{\text{RIVM}}(t_0)$, minus the reported infections one infectious period ago $I_i^{\text{RIVM}}(t_0 - 1/\gamma)$. We assume that only a fraction a of the total number of currently infected individuals gets detected by the RIVM. A fraction $(1-a)$ is not tested by RIVM and unknown to RIVM. We initialize as follows:

Expressed in number of individuals, we have

$$I_x(t_0) = \frac{1}{a}\left(I_x^{\text{RIVM}}(t_0) - I_x^{\text{RIVM}}(t_0 - 1/\gamma)\right), \qquad R_x(t_0) = \frac{1}{a} I_x^{\text{RIVM}}(t_0 - 1/\gamma).$$

Exposed people are people who become infectious after the latent period has ended. The number of people exposed at time t_0 is equal to the infections one latency period from now $1/aI_x(t_0 + 1/\nu)$, minus the infections at time t_0, $1/aI_x(t_0)$. In terms of reported cases by RIVM, we can write

$$E_x(t_0) = \frac{1}{a}\left(I_x^{\text{RIVM}}(t_0 + 1/\nu) - I_x^{\text{RIVM}}(t_0)\right).$$

Then, lastly, by using the conservation law, we initialize $S_x(t)$ as

$$S_x(t_0) = N_x - E_x(t_0) - I_x(t_0) - R_x(t_0).$$

For concentration values, we divide by the subpopulation sizes:

$$s_x(t_0) = \frac{S_x(t_0)}{N_x}, \qquad e_x(t_0) = \frac{E_x(t_0)}{N_x}, \qquad i_x(t_0) = \frac{I_x(t_0)}{N_x}, \qquad r_x(t_0) = \frac{R_x(t_0)}{N_x}.$$

This initial state can be used as input for the adaptive mobility regions or as a starting point for evaluating the performance of different divisions.

F.2 Concentration measure

To quantify the concentration of the infections, we compare the distribution of the infections over the municipalities to the distribution of the population over the municipalities. For this, we introduce an entropy-type measure. Let p_i^{inf} denote the fraction of infectious individuals that live in municipality i and let p_i^{pop} denote the fraction of the population that lives in municipality i. If the infections would be distributed evenly, then these two distributions would be equal. We measure the concentration of the infections by $1 - \mathrm{e}^{-D_{\text{KL}}}$, where D_{KL} denotes the Kullback-Leibler divergence from the infection distribution $\left(p_i^{\text{inf}}\right)_{i\in\mathcal{A}}$ to the population distribution $\left(p_i^{\text{pop}}\right)_{i\in\mathcal{A}}$.

Note that this measure does not depend on the total number of infections. When each municipality has the same percentage of infected individuals, the concentration will equal 0. On the

other hand, when all infections are located in the same municipality, then it will be close to 1.

This allows us to quantify the concentration of the infections for our initializations on a scale from 0 to 1. Synthetic initializations are close to 0 and 1, respectively, while historical initializations have intermediate values. The concentration values of initializations based on the historical data obtained from the Dutch National Institute for Public Health and the Environment (RIVM) are shown in Figure 5.4. At the beginning of the epidemic, there were a few local outbreaks, resulting in a high concentration. As the outbreak became more widespread, the concentration went down, reaching a minimum at the beginning of April. The period after April is characterized by small local outbreaks, leading again to higher concentration values.

Appendix G

Optimization Heuristics

In Section 5.2.4, we argued that finding the optimal division that optimizes Equation (5.8) becomes computational infeasible as the number of different divisions can be very large; see Equation (5.9). To improve on this, we define coarser sub-divisions as starting points for the optimization procedure. Instead of letting Louvain operate at the level of single municipalities, we will start from an initial division of the municipalities into sub-regions that will not be further divided by the algorithm. This way, Louvain is guaranteed to result in a division that cannot be improved upon by merging two regions or by moving one sub-region to another region [77].

There are multiple ways to choose such initial divisions. One can use existing administrative divisions such as the twelve provinces of the Netherlands or its twenty-five security regions. Such administrative divisions may already be known by the public, making it easier to enforce restrictions based on them. A disadvantage can be that such regions have not been defined with an outbreak of an infectious disease in mind and therefore do not pose natural boundaries to transmission. Neither have they been defined on the basis of economic activity. We can also use other criteria to find other initial divisions, for example based on behavior of individuals that relates to transmission or to economic activity. Human mobility may be a good indicator for both of these. We give two criteria for obtaining such initial divisions.

G.1 Mobility regions.

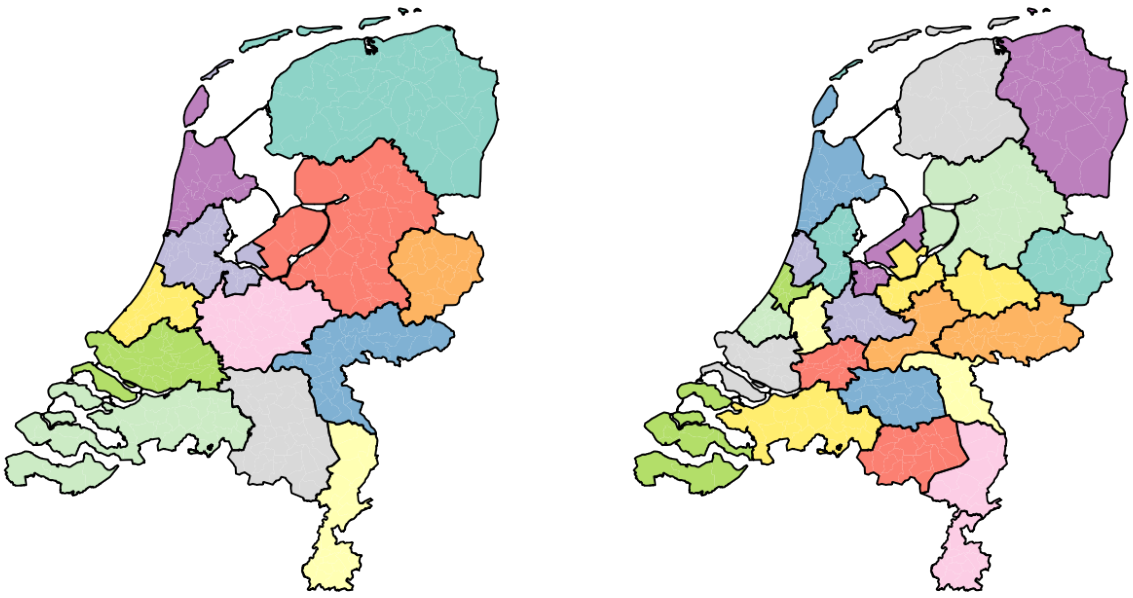
In network science, the objective of community detection is to partition the nodes of a network into groups that are more highly connected internally than externally [78]. Community detection has been applied to mobility networks [79, 80], resulting in divisions into regions that are coherent with respect to mobility.

Currently, the most popular community detection method is to optimize a quantity called *modularity*, which computes the weight inside the communities minus the expected weight for a random network with the same weights. For mobility data of the kind that we rely on, the modularity of a division $\mathcal{D} = \{D_1, \ldots, D_{|\mathcal{D}|}\}$ is given by

$$\text{Modularity}_\eta(\mathcal{D}) = \frac{1}{M(\mathcal{A},\mathcal{A})} \sum_{D\in\mathcal{D}} \left[M(D,D) - \eta \frac{M(D,\mathcal{A})\, M(\mathcal{A},D)}{M(\mathcal{A},\mathcal{A})} \right],$$

where $M(A,B) = \sum_{x,y\in A} M_{xy}$ is the mobility between the sub-populations in $A \subseteq \mathcal{A}$ and in $B \subseteq \mathcal{A}$, and η is the resolution parameter, which controls the granularity of the found communities [81]. Larger values of η result in divisions consisting of more regions.

We iterate the Louvain algorithm [82] to optimize modularity. We will refer to regions obtained by this optimization as *mobility regions*. We note that the number of mobility regions that result from applying this method for a given resolution parameter is, in general, not known beforehand. Therefore, we apply this method for a variety of resolution parameter values to obtain divisions consisting of different numbers of mobility regions. For example, by trial and error, it was found that the choice $\eta = 2$ resulted in a division into 12 mobility regions, so that this division has comparable granularity as the division into the 12 provinces. Figure G.1 shows examples of our mobility regions for two values of the resolution parameter.



(a) Resolution parameter $\eta = 2$

(b) Resolution parameter $\eta = 4$

FIGURE G.1: Two example outcomes of our mobility region heuristic for two different resolution parameters.

A comparison of these two divisions based on mobility is shown in Figure 5.5. We see that these mobility regions are indeed more coherent in terms of mobility. In particular, the city of Almere is in the same mobility region as the city of Amsterdam, but in a different province. Mobility regions may reflect the economic and non-local transmission activity of the citizens in a better way. However, they cannot be tailored to the status of the epidemic.

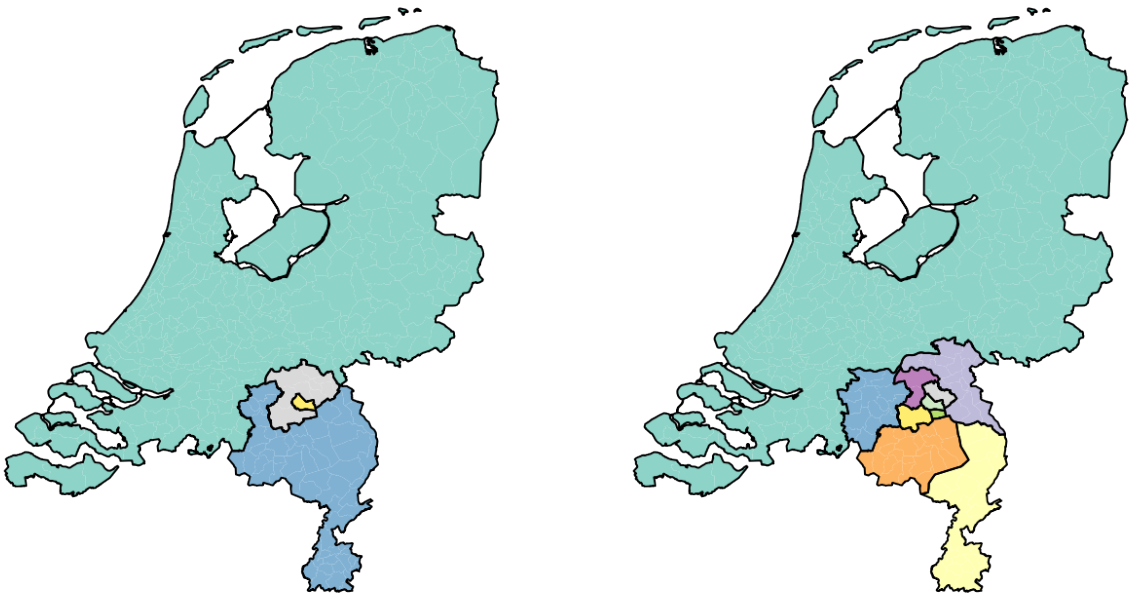
G.2 Adaptive mobility regions.

The running time of our optimization algorithm depends heavily on the number of sub-regions in our starting division. Note that for the mobility regions, the resolution parameter controls the resolution of the division globally. However, high resolution is mostly needed around locations where a lot of infections occur. We next introduce a modification to the modularity function of the previous section to obtain initial divisions that have a higher resolution near such critical locations. Infections are due to meetings between infectious and susceptible individuals. Therefore, a good heuristic would aim for a division that separates infectious individuals from susceptible ones as much as possible. For a region $D \subset \mathcal{A}$, let $I(D), S(D)$ denote the number of infectious and susceptible individuals living in D, respectively. We maximize the

function

$$\text{AdaptiveModularity}_{\zeta}\left(\mathcal{D}\right) = \sum_{D \in \mathcal{D}} \left[M\left(D, D\right) - \zeta \frac{I\left(D\right) S\left(D\right)}{N} \right],$$

where ζ is a resolution parameter that, similarly to the η parameter of modularity, determines the granularity of the obtained division. In particular, by varying ζ, the method results in divisions with different numbers of regions. Despite the fact that ζ has a similar role to the η of the mobility regions, it does have a different dimension and differs an order of magnitude. Again, the Louvain algorithm is iterated for optimization. We call the resulting regions *adaptive mobility regions* because the resolution is locally adapted to the state of the epidemic. Figure 5.6 compares a division into mobility regions to a division into adaptive mobility regions for the concentrated initialization. We see that the adaptive mobility regions indeed have a higher resolution around the critical area of the superspreader event in Uden. Figure G.2 shows examples of our adaptive mobility regions for two values of the resolution parameter.



(a) Resolution parameter $\zeta = 8000$

(b) Resolution parameter $\zeta = 30000$

FIGURE G.2: Two example outcomes of our adaptive mobility region heuristic for two different resolution parameters. The model is initialized following a concentrated initialization of 1000 individuals in the municipality of Uden, see Appendix F. Uden colored yellow in the left figure.

Appendix H

Robustness of Chapter 5 results

Here we discuss stochastic effects on our SEIR model. In Section H.2 we present a sensitivity analysis for the results of Section 5.4.

H.1 Stochastic effects

In Chapter 5, we work with a *deterministic* SEIR model. In this model, the spread of the infection is modeled using a discrete-time evolution model. This makes that numbers are no longer integers. As a result, one obtains the well-known anomaly that when there would be 10^{-6}, say, 'people' still infected, the infection would not die out and could grow exponentially again when $\mathcal{R}_{\text{eff}} > 1$.

Simulations show that stochastic rounding has a minor effect when the numbers of infections are large, see Figure H.2. However, in the case where an initialization is highly concentrated, the spread to neighboring regions will be strongly affected by stochastic rounding, as we might expect. In an extreme case where only 1 person is initially exposed, stochastic rounding can deviate far from the deterministic outcome, and also different simulations can deviate substantially from each other. Figure H.3 and Figure H.4 show results for different simulations with a time horizon of 30 days and 100 days, respectively, with one initial infected individual. There is a positive probability of the infection dying out, while if the infection does not die out, it grows due to $\mathcal{R}_{\text{eff}} > 1$.

As expected, the influence of stochastic effects is relatively large for small numbers of infections. In Section5.4, we initialize at least 1000, comparable to Figure H.2, where the effect of stochastic simulation is relatively small.

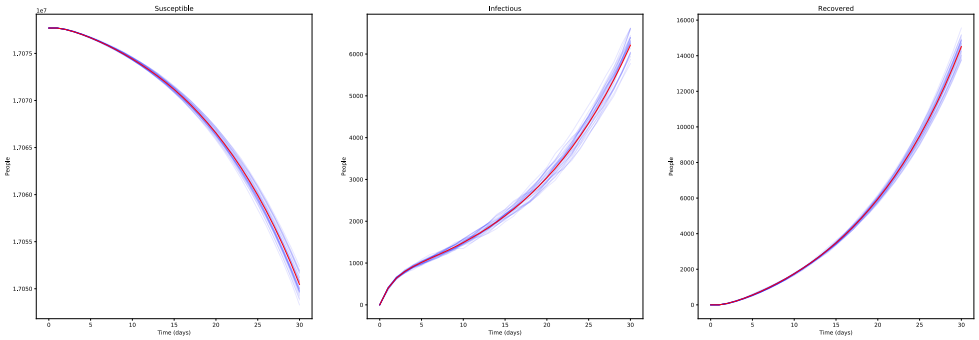


FIGURE H.1: 2000 individuals are initially exposed in the municipality of Uden. The total number of people in the Netherlands who are susceptible, infectious, and recovered are presented for a period of 30 days. Model parameters can be found in Table 5.1 Deterministic outcome is in red, stochastic rounding simulations are in blue.

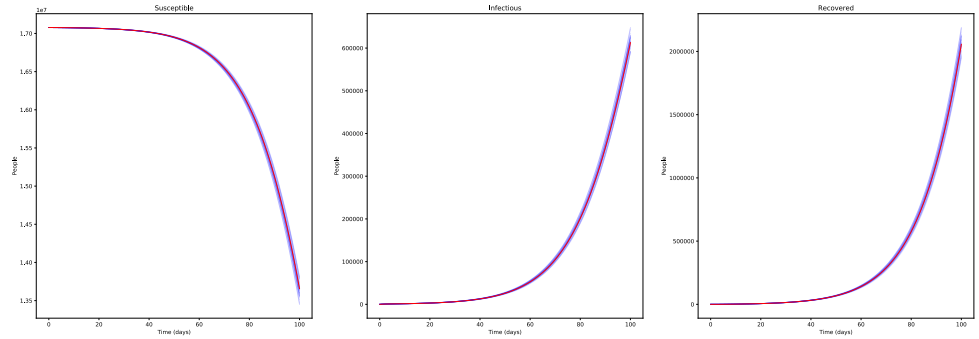


FIGURE H.2: 2000 individuals are initially exposed in the municipality of Uden. The total number of people in the Netherlands who are susceptible, infectious, and recovered are presented for a period of 100 days. Model parameters can be found in Table 5.1 Deterministic outcome is in red, stochastic rounding simulations are in blue.

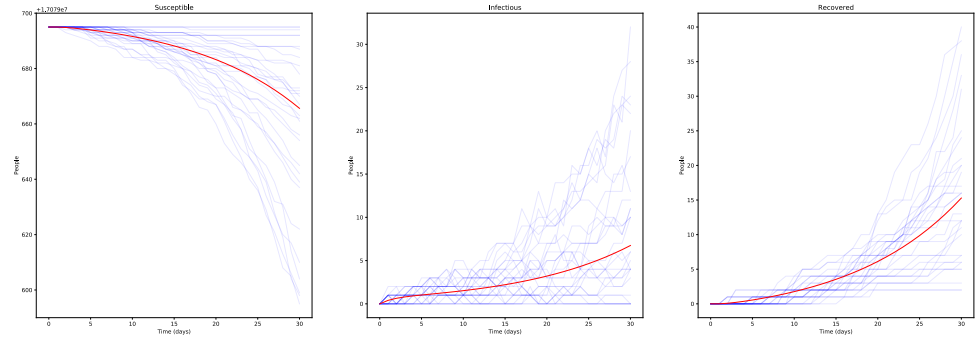


FIGURE H.3: One individual is initially exposed in the municipality of Uden. The total number of people in the Netherlands who are susceptible, infectious, and recovered are presented for a period of 30 days. Model parameters can be found in Table 5.1 Deterministic outcome is in red, stochastic rounding simulations are in blue.

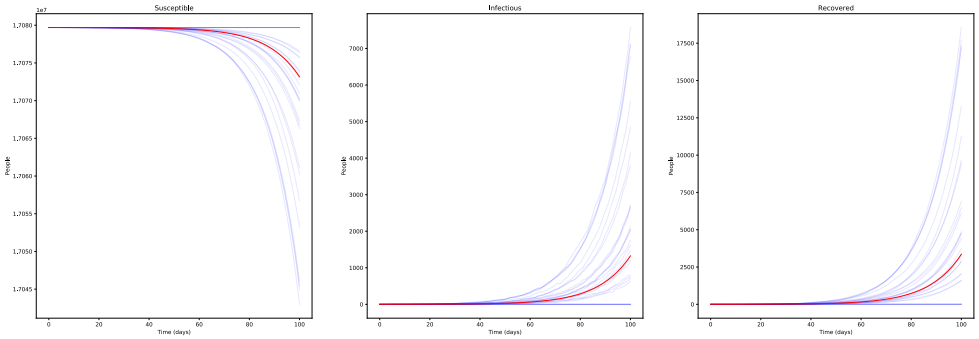
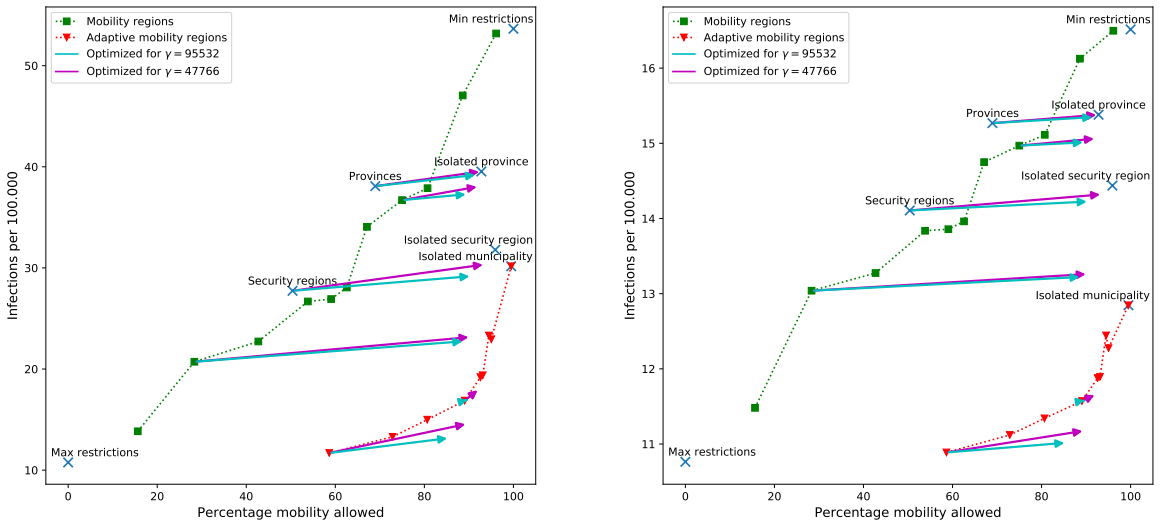


FIGURE H.4: One individual is initially exposed in the municipality of Uden. The total number of people in the Netherlands who are susceptible, infectious, and recovered are presented for a period of 100 days. Model parameters can be found in Table 5.1 Deterministic outcome is in red, stochastic rounding simulations are in blue

H.2 Sensitivity Analysis

In this section, we briefly show the sensitivity of our results with respect to the different input parameters. We analyze the sensitivity of the results to the model parameters by varying parameters one-at-a-time, using the values from Table H.1. The results in Figures 5.10(a), 5.10(b), 5.11(a), 5.11(b) are recalculated with different parameters of the model and are shown in Figures H.5, H.6, H.7, H.7, H.8(b), and H.9.

Differences are most noticeable for simulations where the locations of the infectious people are concentrated. But even then, our conclusions are fairly robust against changing parameters. Infection counts on the y-axis do differ a lot, but the shape of the curves stay roughly the same, and thus, our conclusions also remain unchanged. This effect is clearly visible when we change the effective reproduction number to 1 and to 2.5. Then, infections after 30 days are at 9000 or 50,000, but the relative performance of different divisions is similar.

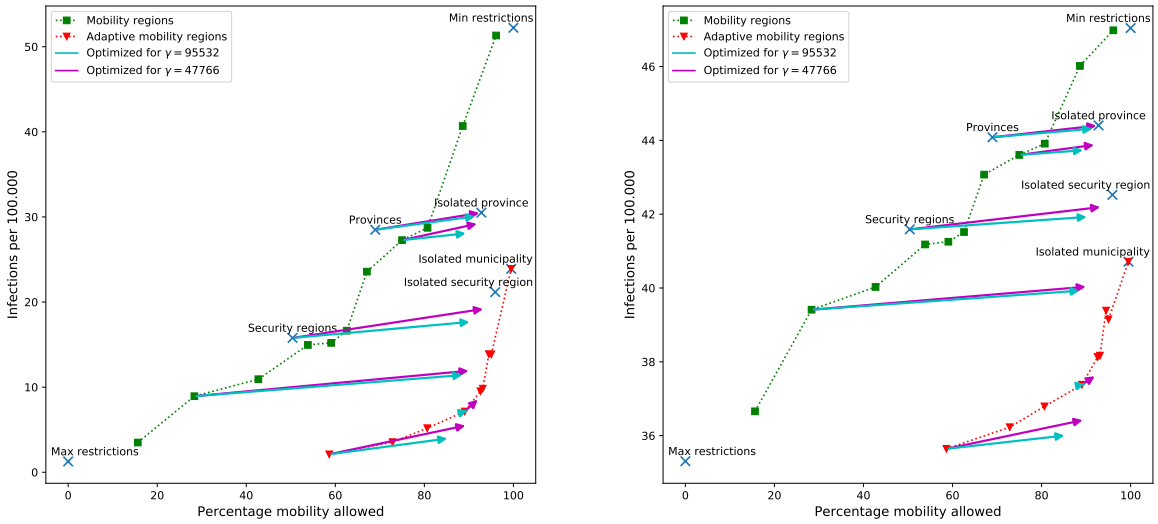


(a) $a = 0.01$ (b) $a = 1$

FIGURE H.5: Analyzing robustness of Figure 5.10(b) with respect to the fraction of tested people.

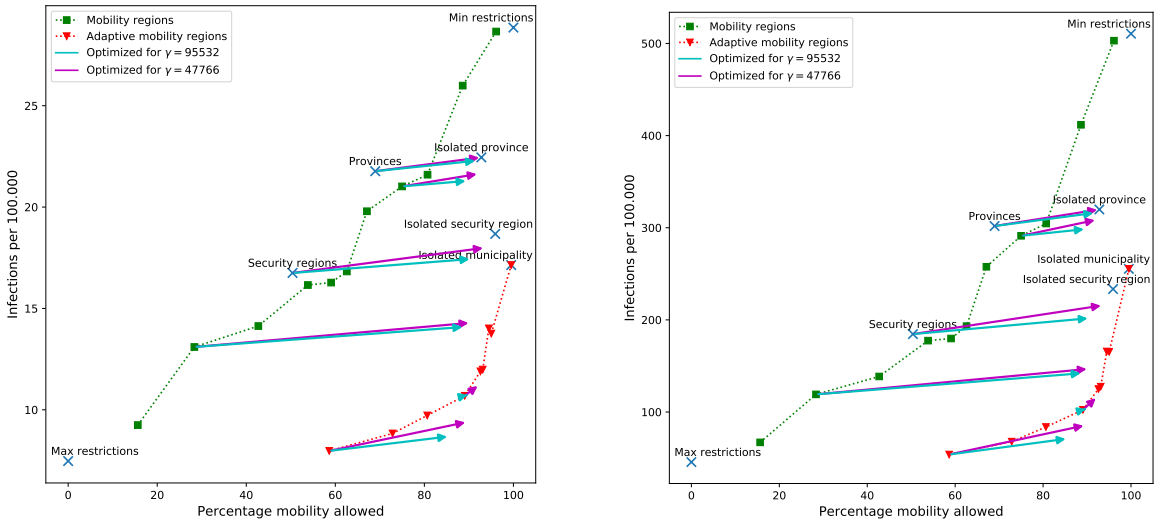
Name	Variable	Base value	Lower value	Upper value
Fraction tested	a	$1/15$	0.01	1
Fraction local contacts	p	0.5	0.1	0.9
Effective reproduction number	$\mathcal{R}_{\text{eff}}$	1.25	1	2.5
Infectious period	$1/\gamma$	5	3	7.5
Latent period	$1/\nu$	5	3	7.5

TABLE H.1: Ranges of parameters for our sensitivity analysis



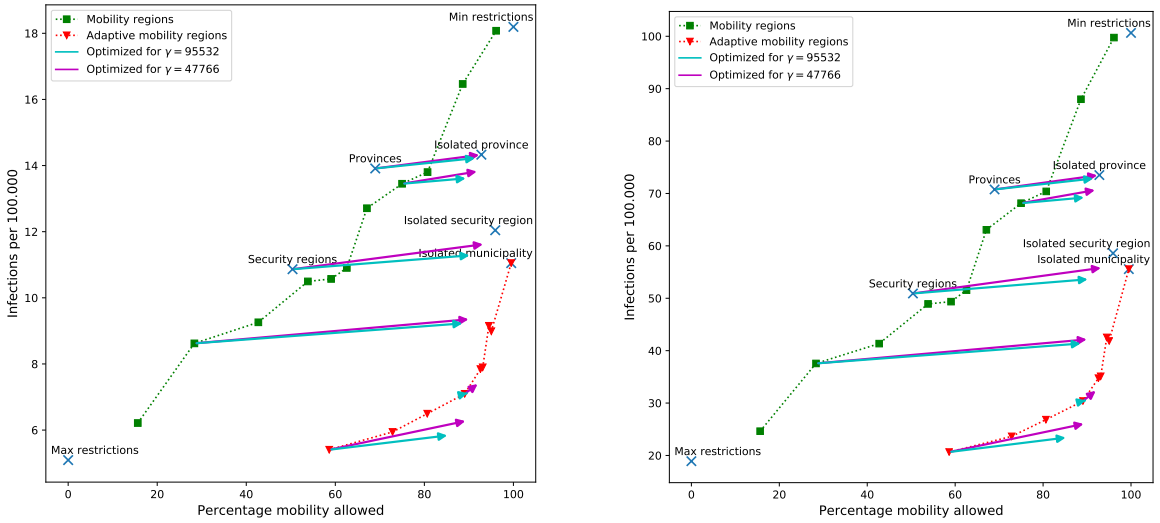
(a) $p = 0.1$ (b) $p = 0.9$

FIGURE H.6: Analyzing robustness of Figure 5.10(b) with respect to the fraction of local contacts.



(a) $\mathcal{R}_0 = 1$ (b) $\mathcal{R}_0 = 2.5$

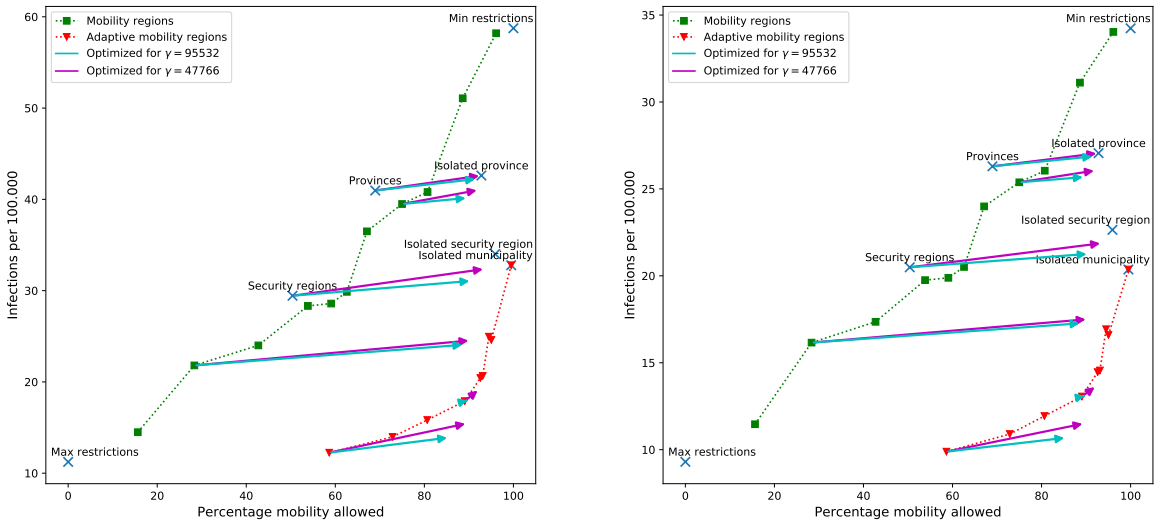
FIGURE H.7: Analyzing robustness of Figure 5.10(b) with respect to the effective reproduction number.



(a) $\omega = 3$ days

(b) $\omega = 7.5$ days

FIGURE H.8: Analyzing robustness of Figure 5.10(b) with respect to the infectious period.



(a) $\nu = 3$ days

(b) $\nu = 7.5$ days

FIGURE H.9: Analyzing robustness of Figure 5.10(b) with respect to the latent period.