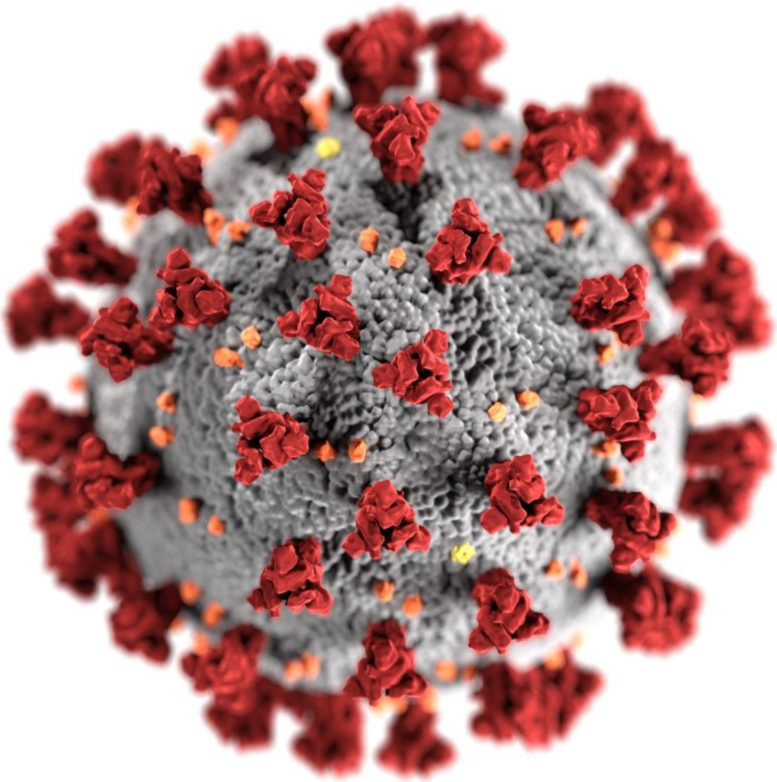
A simulation study to the prediction of COVID-19 in the Netherlands

Master thesis
February, 2021



A project made possible by a contribution of the BMS COVID-19 fund

Author:
C.J. Tenhagen (Cyrelle)

Supervisors of University of Twente:
Dr. E. Topan
Dr. C.G.M. Groothuis – Oudshoorn

Educational program:
MSc. Industrial Engineering & Management
Specialization: Production and Logistics Management
Orientation: Operations Management in Healthcare

UNIVERSITY OF TWENTE.

Cover photo: (Sciensano, 2020)

Preface

Dear reader,

Hereby the final version of my master thesis, which serves as the end of my master Industrial Engineering and Management at the University of Twente. This project was made possible by a contribution of the BMS COVID-19 fund. Time flew by quickly, but I definitely had a great time studying in Enschede. I got to know many amazing people and learnt a lot, especially during the last part of my master studies. Finally I think I discovered where my interests lie.

First, I want to thank Engin Topan and Karin Groothuis-Oudshoorn for giving me the opportunity to conduct my thesis at the University of Twente. Their guidance during my graduation helped me a lot and I especially appreciate the relation we built in this short amount of time. I consider myself very lucky for the opportunity to finalize my studies with them.

Second, I want to thank my boyfriend, rowing mates, and family for getting me through this strange pandemic period. Training with my rowing mates helped me to stay motivated throughout my whole graduation period. Unfortunately graduation also means my last year of rowing, but I had an amazing time and definitely learnt a lot. Yet I am also looking forward to the new challenges I am going to face.

I hope you enjoy reading my thesis!

Cyrelle Tenhagen

Enschede, February 2021

Management summary

This project was made possible by a contribution of the BMS COVID-19 fund from the University of Twente. With this study, we gained understanding in the spread of COVID-19 in the Netherlands. The study contributes with a system dynamics model that predicts the spread of the pandemic and with insights developed in measures and their impacts.

The pandemic of COVID-19 affects and has been affecting the social and economic life of the world severely. Governments all over the world are taking several measures to control the pandemic. Yet, governments do not know exactly how much impact measures have on limiting the spread of the pandemic, and how factors such as weather impact spread. Our model predicting spread of COVID-19 includes key factors and measures with a reasonable influence on the spread. Factors we study are incubation period, infectious period, asymptomatic fraction, reproduction rate, fatality ratio, age, weather (temperature, humidity, wind speed), contact rate (population density, adoption of government measures, places of infection), testing capacity. Measures we study are event allowance, school openings, catering services openings, facemasks, and self-quarantine.

A way to express spread of a virus is with the effective reproduction rate, indicating the rate of transmission. The effective reproduction rate is normally determined with help of the number of infected cases or the number of confirmed cases on a day, depending on availability of data. A drawback of this approach is that the effective reproduction can only be accurately determined two weeks after, since the incubation period, testing delay and reporting delay take time. With backward linear regression, we develop a regression model, called $R_{e\ linear}(t)$, that can predict the effective reproduction rate on day *t* with help of values for key factors and measures on day *t*.

$$\begin{aligned} R_{e\ linear}(t) \ = \ & 1.22 - 0.013 * \text{average temperature} - 4.179 * \text{staying home behaviour} + 0.578 * \text{traveling behaviour} \\ & + 0.066 * \text{school openings} + 0.021 * \text{catering service openings} - 0.109 * \text{event allowance} + 0.095 \\ & * \text{facemasks} \end{aligned}$$

In this model, adoption of government measures, expressed with the variables staying home behaviour and traveling behaviour, has major impact on the effective reproduction rate. Ambient temperature also shows a considerable impact, with a higher reproduction rate when temperature decreases. Of the measures we study, event allowance has most impact. Impact of this measure is comparable to the impact of traveling behaviour. School openings, catering service openings, and facemasks, all show a positive relation to the effective reproduction rate. Consequently, spread increases with stricter measures, which is remarkable because this indicates that stricter measures lead to more spread. An explanation for these positive relations could be that the effect of a certain measure (e.g. school openings) is partially expressed with another variable (e.g. traveling behaviour), since traveling behaviour decreases when schools are closed. These kind of interactions between variables can be expressed with interaction terms. We check whether interacting terms add value for prediction of spread, by comparing performance of this interaction model with the linear model. We express performance of the models in test error, indicated with Mean Squared Error (MSE), and R^2. An R^2 close to 1 indicates that the model explains a large proportion of variance in the response variable. While the interaction model outperforms $R_{e\ linear}(t)$ in performance, we consider the model to be overfit. This means that patterns found with training data do not exist in unseen data.

In our system dynamics model, we determine the number of infected cases with $R_{e\ linear}(t)$. Next to the number of infected cases, we have to predict the number of confirmed cases on a day. We estimate the number of confirmed cases with help of multiple linear regression. The number of confirmed cases on day *t* is estimated with the number of tests and the number of infected cases on day *t*. We found that a logarithmic response improves prediction of the number of confirmed cases, leading to the following regression model for the number of confirmed cases.

$$confirmed\ cases\ (t)\ =\ e^{4.86+4.74\mathrm{e}-05*\mathrm{num}ber\ of\ tests+1.329\mathrm{e}-04*infected\ cases}$$

With help of the regression models and the system dynamics model, we determine the effect of key factors and measures on spread simultaneously. We propose several policies to prevent spread of COVID-19. Performance of these policies is compared to performance of the actual policy, in terms of spread and number of days with strict measures. Two of the proposed policies, elimination of the virus and doing nothing to prevent spread, are not considered to be feasible in the case of COVID-19, since these policies would cause major damage to public health and economy. Feasible policies we study are mitigation, curbing, testing capacity (high/low), and facemasks (implemented/not implemented). Of these policies, mitigation and curbing determine their policy based on signal values. Where mitigation accepts circulation of the virus to a certain extent, curbing strives for little infections as possible. The Dutch government developed a route map to indicate which measures to implement when certain signal values are reached. For curbing we consider two additional versions, referred to as curbing type 2 and curbing type 3, to see how stricter measures affect spread. In short, curbing type 1 acts according to signal values for the number of confirmed cases per day in the route map of the government, curbing type 2 applies only very soft measures or very strict measures, and curbing type 3 applies, regardless of spread, two weeks of soft measures and two weeks of strict measures sequentially.

The policy with high testing capacity shows the lowest values of spread of all feasible policies. Curbing type 2 and type 3 show the lowest values of spread thereafter, yet with a considerably lower number of days with strict measures than the policy with high testing capacity and the actual policy. Mitigation results in lower values of spread than actual values on 31 November, but at the same time needs a much higher number of days with strict measures than the actual policy. Curbing type 1 results in somewhat higher values of spread than the actual policy and somewhat lower number of days with strict measures. We question the reliability of results for the policy facemasks. According to the model, spread would greatly reduce when no facemasks would have been implemented and would greatly increase when they are implemented. In reality, facemasks were obliged from 1 December until date of writing, and spread has not greatly reduced or increased within this time. Therefore we think we might not have sufficient observations for this measure to determine its effect, or this measure interacts with other factors or measures. The reason for curbing type 2, curbing type 3 and high testing capacity to perform well is found to be the high number of tests per day, that leads to a relatively high number of confirmed cases in quarantine. Curbing type 2 and type 3 are also found to perform well since these policies quickly react to spread by applying only strict measures. Curbing type 2 was found to perform better than curbing type 3 due to the timing of implemented measures.

We study robustness of our model with multivariate and univariate sensitivity analysis, where we change input values for sensitivity parameters. Sensitive parameters are considered to be the basic reproduction rate, initial number of infected cases (on 2 February), quarantine fraction (number of confirmed cases effectively entering self-quarantine), and the fraction of asymptomatic cases. We learn from the sensitivity analysis that our model is very sensitive for changes. For example, the pandemic would have stopped existing after the first peak with a basic reproduction rate of 2. Therefore, calibrated values of input parameters in our model are considered to be good approximations, but might deviate a little in reality. We conclude from the sensitivity analysis that it is very important to combine a high testing capacity with a high quarantine fraction. High testing performs very well in terms of spread, yet when the quarantine fraction becomes low, curbing type 2, curbing type 3 and mitigation outperform the policy high testing. When the quarantine fraction becomes high, policies with actual, high, and low testing capacity all outperform remaining policies in terms of spread. Concluding, combining a high testing capacity with a high quarantine fraction is even more important to reduce spread than strictness and timing of implementation of measures.

Table of Contents

1. Introduction

On 31 December 2019, the Country Office of the WHO in the People's Republic of China was informed about several cases of "viral pneumonia" with unknown cause. The viral pneumonia outbreak was identified as a novel coronavirus a few days later. On 30 January 2020, the WHO declared the outbreak of the coronavirus to be of international threat for public health and on 11 March 2020 the WHO declared the disease as a pandemic (WHO, 2020b). This novel coronavirus is called SARS-CoV-2 (Severe Acute Respiratory Syndrome coronavirus-2). The disease that is caused by SARS-CoV-2 is called COVID-19 (Coronavirus Disease 2019). The pandemic has been affecting and still affects the social and economic life of the world severely. Accurate prediction of the spread of the virus is key to deal with the pandemic and keep our social and economic life going. In this chapter, the characteristics and development of the virus will be discussed with a focus on the Netherlands. At the end of the chapter, the research motivation, research objective and research approach of this study are provided.

1.1 Development and characteristics of a novel coronavirus

1.1.1 Virus transmission

The development of effective public health and infection prevention measures are essential to reduce the transmission rate of SARS-CoV-2. It is however not yet fully understood how, when and in what types of settings the virus spreads. Two important factors that have to be understanded to reduce spread are knowing how the virus is spreading and when an infected person can spread the virus. For both of these factors is not (yet) enough evidence available to be able to apply effective measures. According to the WHO, current evidence is suggesting that the virus is transmitted mainly when an infected person is in close contact with another person. Virus spreads primarily via liquid particles that can be different in size. Smaller particles are called "aerosols" and larger particles are called "respiratory droplets". Respiratory droplets are more commonly causing spread, according to current evidence (WHO, 2020d). Infected persons spread respiratory droplets by talking, coughing, sneezing, or singing. Evidence further shows that SARS-CoV-2 can survive outside the body only for a limited amount of time, ranging from hours to days. This depends on type of surface, temperature, and the humidity of the environment. The WHO states that high-quality research is urgent for understanding the relative importance of different routes of transmission. Most important according to the WHO is to find out the role of airborne transmissions without aerosol generating procedures, risk factors and settings of superspreading events, the required dose of the virus needed for transmission, and the extent of transmission both pre-symptomatically and asymptomatically (WHO, 2020c).

1.1.2 Symptoms

The *incubation period* of COVID-19 is estimated to be on average five to six days, but can also be as long as 14 days. After this period, main symptoms of the disease are fever, cough, fatigue, slight dyspnoea, sore throat, headache, conjunctivitis, and gastrointestinal issues. A mild form of the disease is shown in about 80-90% of the cases. Serious symptoms are shown in approximately 10% of the cases and a critical condition develops in around 5% of the cases. These critical cases show pneumonia, shock, respiratory failure, multiorgan failure, and in the worst case death. Risk groups for a poor outcome of the disease mainly include higher aged individuals, or individuals with ischaemic heart disease, diabetes mellitus, hypertension, and chronic lung disease (Pascarella et al., 2020). Next to cases with symptomatic disease, there are also cases with an asymptomatic disease. The *fraction of asymptomatic cases* is not yet discovered. According to Katri Manninen, this fraction is estimated to be 40%. Katri Manninen developed a graph that visualizes the typical progress of COVID-19, with estimates updated in October 2020 (see Figure 1). This graph is made according to estimates from

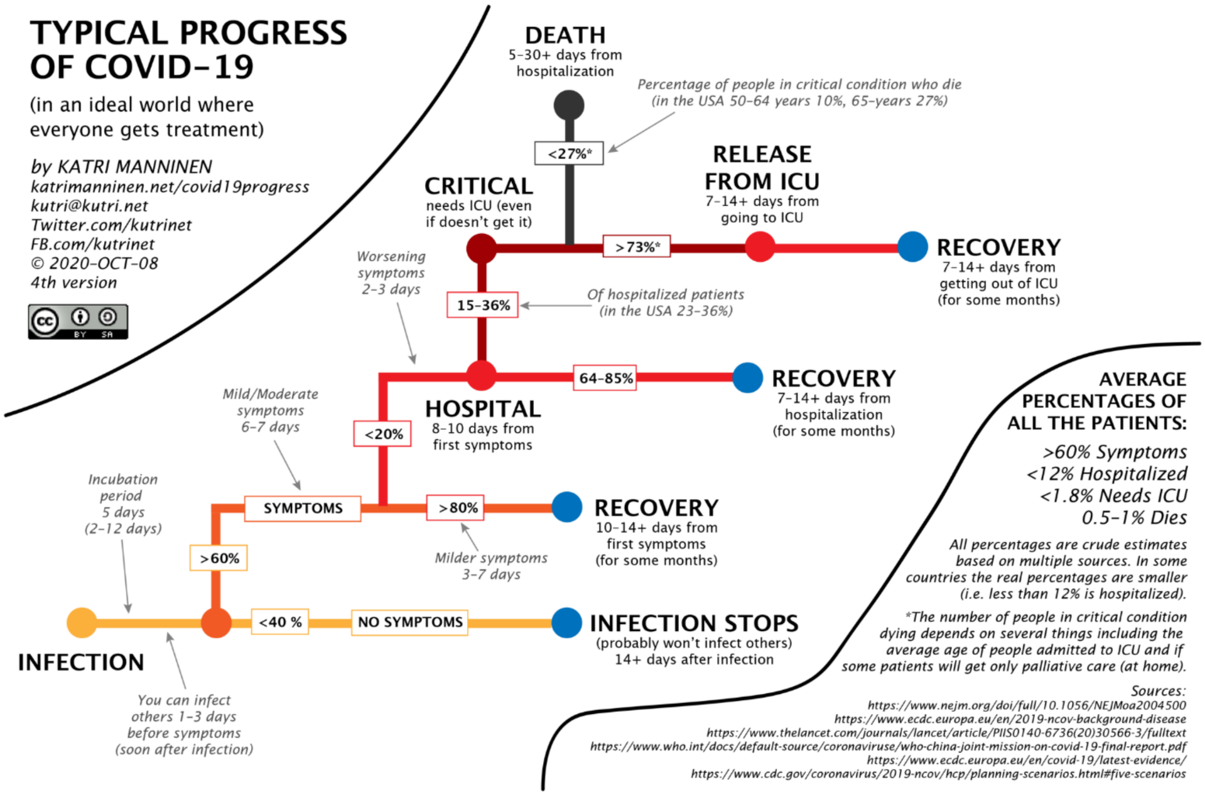


Figure 1 Development of the typical progress of COVID-19

among others the WHO, ECDC (European Centre for Disease Prevention and Control), and the CDC (Centers for Disease Control and Prevention). What we see in this graph is that of all infected cases approximately 0.5-1% dies (Katri Manninen, 2020). The approximation is lower than some previous coronaviruses, with SARS-CoV and MERS-CoV having a *mortality rate* of respectively 10% and 35% (Pascarella et al., 2020).

1.1.3 Worldwide spread of the virus

At the beginning of March 2020, the number of new cases in the world started to grow exponentially. Approximately one month later, the number of cases per day flattened. Flattening of the curve was the result of social distancing measures that were implemented everywhere in the world. Many countries implemented strict distancing measures (e.g. *lockdown)* to be able to handle the number of people with the disease. After a while, economic and social pressures faced by governments and organizations forced them to gradually and safely release the social distancing measures again. On date of writing, 19 January 2021, the WHO reported 93,956,883 confirmed cases of COVID-19 worldwide, including 2,029,084 deaths (WHO, 2020e).

1.1.4 Spread of the virus in the Netherlands

On 27 February, the first infected case of COVID-19 was identified in the Netherlands (RIVM, 2020h). This case and the new confirmed cases the few days thereafter were most probably infected in Italy (RIVM, 2020i, 2020g). The number of confirmed cases per day in the Netherlands started an exponential increase around the beginning of March, which we visualise in Figure 2 (RIVM, 2020f). After a peak with approximately 1000 new cases per day (called the "first peak"), the number of confirmed cases started to decline around the middle of April.

At the beginning of May the decline stopped and the number of confirmed cases remained low for a while. Around the middle of July the number of confirmed cases started to increase again.

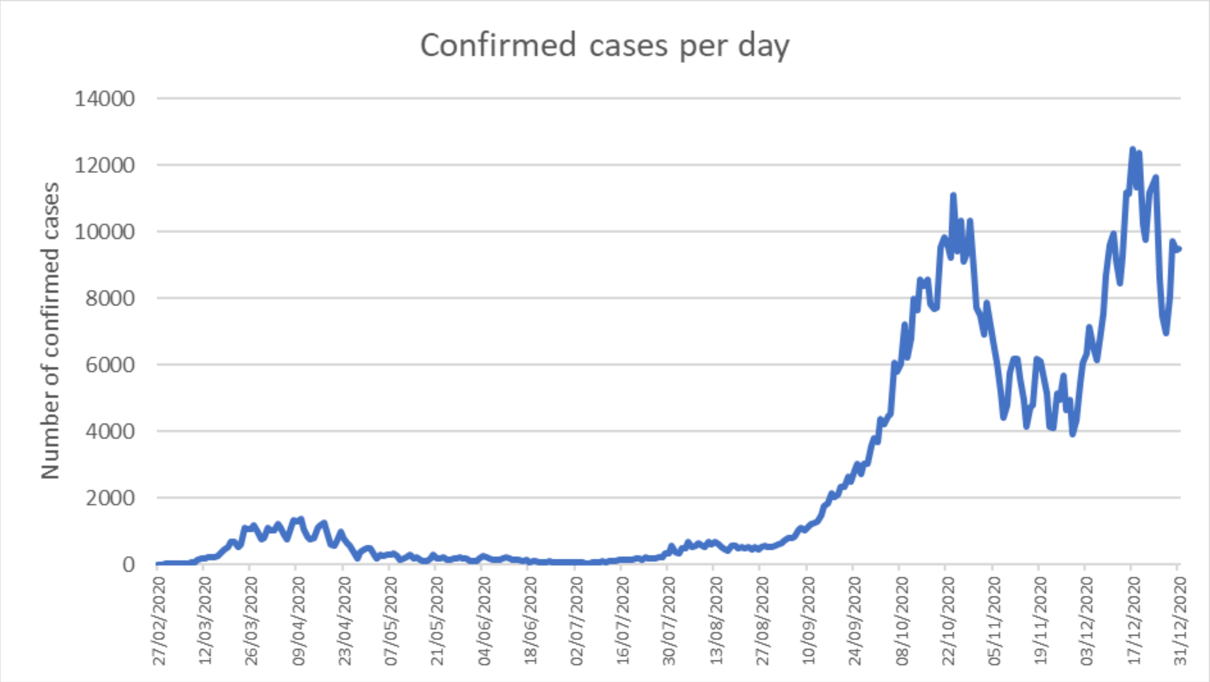


Figure 2 Number of confirmed cases per day in the Netherlands

The first deceased case of COVID-19 in the Netherlands was announced on 6 March 2020. A few days later, a press conference led by the Dutch prime minister was held to express seriousness of the COVID-19 situation in the Netherlands. The prime minister stated that the current economic situation in the Netherlands was good due to a low national debt, low deficits, low unemployment, and promising economic grow expectations. But he mentioned also that the pandemic could have a severe impact on all of this (Rijksoverheid, 2020c).

The decline after the first peak was the result of the implementation of a so-called "intelligent *lockdown*". This lockdown was called intelligent because Dutch residents had somewhat more freedom for movement compared to other countries that implemented a total lockdown (RTLnieuws, 2020b). The lockdown as expected had a severe impact on the Dutch economy and the social life of people. The Dutch government was forced to gradually release the lockdown. There is however lack of scientific evidence on how this can be achieved without causing damage (Block et al., 2020). Limited experience on how to gradually and safely release the lockdown is the cause of a new peak starting in the middle of July. This new peak became problematic after summer holidays, resulting in a significantly higher number of confirmed cases than in the first peak. The number of confirmed cases per day declined a little for a few weeks after implementation of stricter measures, yet increased again at the beginning of December to an even higher number of confirmed cases. At time of writing, the spread of COVID-19 is still worrying and the government is still trying to find the most appropriate measures.

Up until 31 December 2020, the RIVM (the Dutch National Institute for Public Health and Environment) counted a total of 529,304 confirmed cases of COVID-19. The total number of cases that were hospitalized on this date is 27,738 and a total number of confirmed deaths of 11,627 (RIVM, 2020f).

1.2 Research motivation

The number of new cases is still growing worldwide, the Netherlands is dealing with a third peak and the number of hospitalizations and deaths does not seem to decrease anywhere soon. The desire for a solution becomes greater with time, since problems caused by the virus are growing in number and size. Many people deal with psychological distress due to strict measures and the impact on the economy is severe everywhere in the world. However, governments are forced to take these measures to ensure that the number of diseased people remains manageable and hospitals are able to handle all care (Nicola et al., 2020; Qiu et al., 2020).

1.2.1 Problem cluster

To be able to improve impact of COVID-19 on social life and economy and to prevent hospitals from overflowing, finding the core problem of this problem context is key. The core problem can be found with help of a problem cluster. A problem cluster identifies the cause and effect relationships in the context of a problem, which leads to a core problem. Problems we identified in this study can be found in the problem cluster in Figure 3. Arrows indicate the relationship between problems by pointing from cause to effect. The problems without any cause are possible core problems. By improving (one of) these problems, the problems at the top will be improved as well (Heerkens & van Winden, 2017).

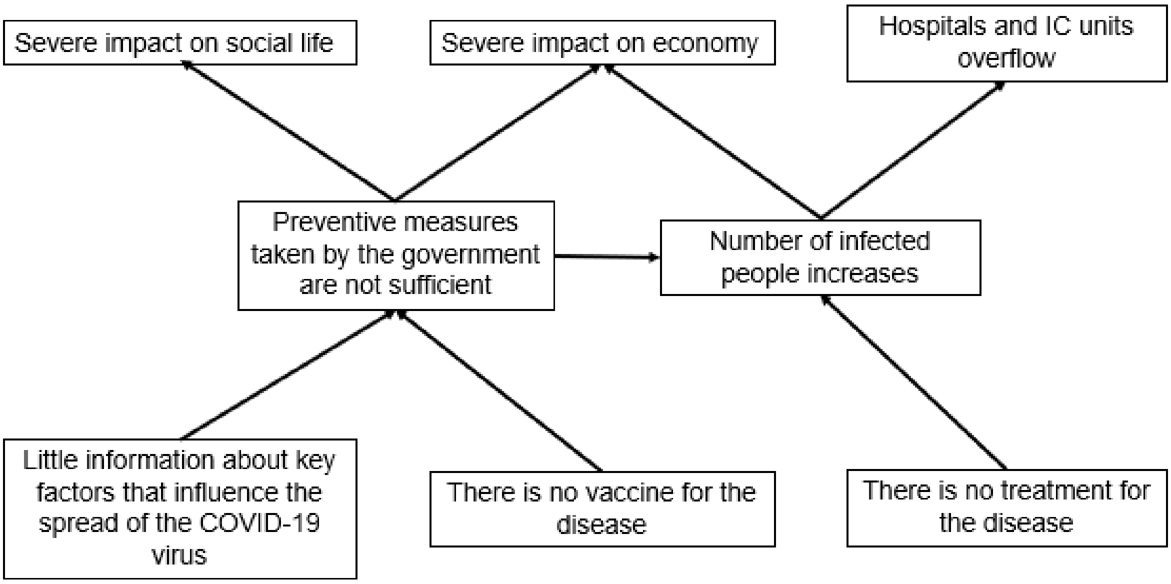


Figure 3 Problem cluster

Little information about key factors that influence the spread of COVID-19

As stated by the WHO, high-quality research is urgent to understand the relative importance of different routes of transmission. This is required to develop effective public health and infection prevention measures, which is essential to reduce the probability of transmitting the virus. Since it is not yet fully understood how, when and in what types of situations the virus spreads, obtaining knowledge about the key factors that influence the spread is important.

Vaccination is not (yet) effective enough to prevent spread of COVID-19

The development of a safe vaccine takes a long time. Normally it takes years to develop a vaccine, but the development of a vaccine for COVID-19 happens rapidly due to the high need for this vaccine worldwide. At the beginning of January 2021, vaccination started in the Netherlands. From this moment, it still takes months before everybody in the Netherlands has received a vaccine and the whole Dutch population is protected against infection.

There is no treatment for the disease

Until now, there is no effective treatment for the disease. Main therapies that are currently used to treat the disease are respiratory therapy, antiviral drugs, and chloroquine/hydroxychloroquine. While there were many therapies proposed to prevent infection or treat the disease, only implementing lockdown measures has shown to be effective for decreasing the rate of transmission. Though global and economic consequences of lockdown measures are severe (Pascarella et al., 2020).

1.2.2 Core problem

The core problem chosen to be solved in this study is “Little information about key factors that influence the spread of COVID-19”. We chose this core problem because the government can only apply effective measures when there is sufficient information about key factors that influence the spread of the COVID-19 virus. With sufficient information, impact of the virus on social life and economy can be reduced, and the number of diseased people can be controlled. Since it currently still takes months before vaccination leads to a protected population, the need for sufficient information is still relevant. To make the scope of the research manageable and discover the effect of country specific measures at the time this thesis is written, the scope of this study is limited to the spread of COVID-19 in the Netherlands.

1.3 Research objective and questions

Until the start of this project, there has been no unified and holistic approach bringing all valuable knowledge about key factors that influence the spread of the virus together and taking into account preventive measures to model the spread of COVID-19 in the Netherlands. Existing models are quite generic and do not include measures that differ per country (e.g. *closure of schools*). This study will look at the impact of government measures (e.g. lockdown) and other factors (e.g. incubation period, *weather*) on the spread of a pandemic in a country. A factor is identified as a key factor if it considerably influences the spread of the virus. We will built a model to predict the future spread of the pandemic by using data from the current spread. Using the model, the impact of government measures as well as other factors on the spread of the virus can be investigated. Indicators (e.g. number of infections) will be used to be able to evaluate the spread. This study focusses on the spread in the Netherlands by including measures and factors that are relevant for the Netherlands in particular. In short, the objective of the study is as follows:

1.3.1 Objective

Understand the spread of COVID-19 by studying key factors affecting the spread of COVID-19 and the impact of measures taken to reduce the spread of COVID-19 in the Netherlands

The key factors that affect spread of COVID-19 will be studied with help of a simulation model. The simulation model helps to develop insight for policymaking. We chose to use system dynamics to build the model. System dynamics is an approach that can help to understand non-linear behaviour of complex systems over time (Marshall et al., n.d.). The spread of COVID-19 can be seen as such a complex system which the whole world is trying to understand.

The research objective leads to the main research question:

1.3.2 Main research question

What are the impact of government measures on the spread of COVID-19 in the Netherlands and how can we learn from this for future outbreaks and pandemics?

We defined several research questions to be able to answer the main research question. The research questions will be answered with help of sub-questions. The research questions and sub-questions are provided below.

1.3.3 Research questions

1. How did the spread of COVID-19 develop in the Netherlands?

Numbers that indicate the spread in the Netherlands are required to find out to what extent the impact of COVID-19 can be reduced with help of the proposed measures at the end of the report. We use historical data from the RIVM to acquire these numbers. Besides, the development of the spread of COVID-19 and its relation to implemented measures has to be known to be able to develop a reliable model. All of this will be outlined in Chapter 2.

a) What indicators can provide a useful indication for the severity of spread?
b) How did the indicators of spread evolve during the pandemic in the Netherlands?
c) What measures were implemented to prevent spread in the Netherlands?
d) What are the differences in spread of the virus between provinces?

2. What factors and measures should be considered in the model?

Next to the required historical data about spread, key factors that influence the spread of the virus have to be discovered. A literature review will be performed to find out which factors possibly influence spread of the virus, and the relevance of those factors. Besides the factors that will be used as input for the model, we have to study measures that can be implemented in the Netherlands. The government currently implemented several measures to control spread of the virus. However, the extent to which applied measures are implemented could be adjusted. For example, when to open schools or whether to wear facemasks. The literature review will be done in Chapter 3.

- a) What factors possibly influence the spread of COVID-19?
- b) What is the relevance of these factors according to current literature?
- c) What measures can be considered to prevent spread in the Netherlands?
- d) What are the effects of these measures according to current literature?
- e) How to model spread of the virus in the population?

3. How to include the key factors and measures in the model?

It is very likely that not all factors identified in the literature review should be included in the model. Some factors may not considerably influence spread of the virus and can be left out. Of the key factors and measures that are influencing spread in the Netherlands it should be clear how to include these as parameters in the system dynamics model. This will be determined using statistical analysis, which we clarify in the Method (Chapter 4).

- a) What are the key factors influencing spread of the virus in the Netherlands?
- b) What is the relation between different key factors?
- c) What is the relation between key factors and measures?
- d) What input values should be used (parameter estimation)?
- e) How can the system dynamics model be validated?

4. What policy can effectively reduce spread of the virus in the Netherlands?

In Chapter 4 we determined how key factors and measures affect spread of COVID-19 in the Netherlands. In Chapter 5, we propose several policies to identify a combination of measures that can effectively reduce spread. The performance of these policies will be determined with help of a system dynamics model. Evaluation of feasible policies will be done in the Conclusion (Chapter 6).

- a) What policies exist to prevent spread of a virus?
- b) Which policies can effectively reduce spread of COVID-19 in the Netherlands?
- c) How do the policies perform (compared to the currently applied policy)?
- d) How robust is the model for changes in parameters (sensitivity analysis)?

5. How can the outcome of the model be used?

When it is clear which factors are key factors for the spread of COVID-19 and what measures should be implemented to be able to control the spread of the virus, we have to discover how this information can be used in practice. This will be done in the Chapter 6 (Conclusion and Discussion).

- a) How can the outcome of the models be used to identify the best policy to prevent spread of COVID-19 in the Netherlands?
- b) In what way can the outcome of the model be used in the fight against spread of COVID-19 worldwide?
- c) To what extent can the outcome of the model be used in future pandemics?

An overview of the content addressed per chapter is provided is Figure 4.

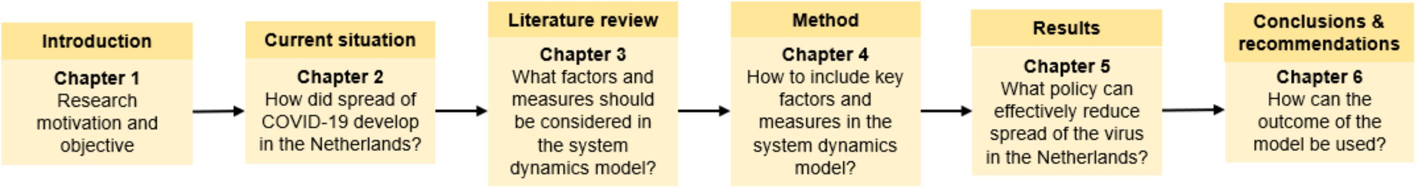


Figure 4 Overview of content addressed per chapter

1.4 Research approach

System dynamics is a simulation modelling method that can capture complex and non-linear relationships between components in a complex system. Being nonlinear models, system dynamics models are sensitive to input parameters. This requires rigorous parameter estimation. Existing estimates and publicly available data will be used for parameter estimation. Statistical learning methods are used to consider possible relations and incorporate the relation between parameters (key factors and measures). Statistical analysis will be done using (multiple, non-) linear regression. Simulation of the model will be done with the system dynamics program Vensim. The system dynamics model will be calibrated to mimic spread and validated by comparing outcomes to actual spread. In the resulting system dynamics model we consider several policies where different measures are implemented. Lastly, a sensitivity analysis will be applied to evaluate the outcome of policies and to determine uncertainty of parameters in our system dynamics model.

Deliverables

At the end of the study, the outcome of the system dynamics model will be used to develop a proposal including appropriate measures for the fight against COVID-19 in the Netherlands.

2. Current situation

To be able to reduce the effect of the pandemic on social and economic lives within the Netherlands, the development of spread of COVID-19 in the Netherlands should be analysed in depth. We do this by describing the development of four indicators of spread between 27 February and 31 December 2020, with a short notion of implemented measures to provide context. We observe varying values of indicators over time in Section 2.2, and varying values across provinces in the Netherlands in Section 2.3. These differences will be analysed briefly.

2.1 Development of spread in the Netherlands

In this section, the development of the most important indicators is outlined. Due to our particular interest and usefulness for predicting the severity of the spread, these indicators include the number of confirmed cases per day (symptomatic and asymptomatic), the number of hospitalized patients, the number of occupied Intensive Care (IC) beds, and the number of deaths per day. All measures we provide below originate from press conferences held by the Dutch prime minister and were gathered from the website of the Dutch Central Government (Rijksoverheid, 2020b). Most measures were communicated to the Dutch residents in the form of advices.

2.1.1 Development of (confirmed) infected cases: Symptomatic

On 27 February, the first infected case was identified in the Netherlands. The number of confirmed cases per day in the Netherlands started an exponential increase around the beginning of March. After a first peak of approximately 1000 new cases per day, the number of new cases per day started to decline from the middle of April. Around 10 May the decline stopped and the number of confirmed cases remains under 200 per day until the middle/end of July. After that, the number of confirmed cases per day increased again. This lead to a second and third peak in October and December.

The first peak

The first mild measures to prevent spread of the virus were implemented on 12 March. Large gatherings were cancelled and people were advised to work from home. Additionally, smaller gatherings (e.g. eating- and drinking occasions and sports clubs) were closed on 15 March and people were urged to keep 1.5 distance. On 23 March some effective measures were refined. All events and gatherings were prohibited and places where physical contact is unavoidable (e.g. hairdressers) were closed. The Dutch government refers to this combination of so-called lockdown measures as an “intelligent” lockdown. This means that people who show symptoms of illness were recommended to stay home. Healthy people were advised to work from home as much as possible. The intelligent lockdown showed to be effective because the number of confirmed cases per day started to increase less rapid and declined after a while. On 9 April, the prime minister spoke positively about the effect of the measures on the corona numbers in a press conference.

Decline after the high peak

On 21 April, a little while after the start of a decline in the new number of cases, the government decided about loosening some measures from 11 May. The first softened measures mostly affected children (e.g. partially opening primary schools). After 11 May, additional measures were gradually being softened. While softening measures, the threat of a possible increase of the number of new cases was taken in mind by the government. Softening of measures was particularly meant for helping the economy and social life. From 11 May, jobs that include physical contact (hairdressers, opticians, etc.) were allowed to open their doors again, while following the 1.5 meters distance rule. Besides softening of measures, the testing capacity gradually increased between 11 May and 1 June. Before 1 June, only persons with a higher risk of a serious course of the disease or hospitalized patients could be tested. After this date, everybody experiencing complaints was able to take a test. This may be the reason for a small increase in the number of confirmed cases around 1 June (see Figure 2).

From 1 June, restaurants, cafes, cinemas, concert halls, and theatres were allowed to open again while following specific preventive measures (e.g. disinfection and keeping distance). From this time, the GGD (Dutch Public Health Services) also started a source- and contact investigation (Dutch: Bron- en Contact Onderzoek; BCO). The goal of this investigation is to prevent further spread of the virus by focussing on identifying persons with whom an infected person has been in (close) contact with (RIVM, 2020j). The number of confirmed cases did not increase after these softened measures, they seemed to decrease even further. Next to opening catering services again on 1 June, public transport started to work their normal schedule again (with 40% capacity). Furthermore, secondary schools opened again with a strict social distancing policy. This means that not all children were able to attend at the same time and schools had to adjust their schedules. From 8 June, primary schools and child care opened fully again (with normal capacity and opening hours) and from 15 June students (MBO, HBO, WO) were allowed to take physical exams and follow practical education again.

At the beginning of June, many Dutch residents started to think about how to spend their summer holidays (Rijksoverheid, 2020f). The government allowed them to spend their holidays outside the Netherlands, but advised to postpone their holidays or to stay in the Netherlands. Around the end of June, the prime minister declared that the most important corona numbers were low or even showed a little decline. For this reason the government decided to soften more measures, even while several (European) countries experienced an increase in number of confirmed cases after softening measures (e.g. Germany)

On 1 July, sports- and fitness clubs, saunas, sports canteens, and casinos were allowed to open their doors again while following strict rules. Outside activities were allowed again with a maximum of 250 people (e.g. soccer stadiums). The government emphasized the need of keeping 1.5 meters distance when stepping into social life like this again. Universities were allowed to fully open again, which was in reality at 1 September due to the summer holidays. In public transport, people were (still) required to wear a facemask but all seats were available again. Night clubs stay closed. We can see in Figure 2 that after 1 July, the number of new cases per day remained under 200 for quite some time after softening of these measures.

Second peak

Approximately one month after the softened measures on 1 July, the number of new cases per day started to increase again (see Figure 2). This time somewhat less exponential than at the beginning of March. Most new cases occurred regional, in big cities like Amsterdam, Rotterdam, Den Haag, and Utrecht. While the number of new cases increased, the government did not decide to implement stricter measures again. After a few weeks, at the beginning of August, spread of the virus started a dangerous rise again. The government decided to take some additional measures. Introduction periods of universities were only allowed when mostly held online and with a strict end time. Catering services had to follow stricter rules and travellers from countries with risk for COVID-19 were obliged to stay in quarantine for two weeks after returning to the Netherlands (even with negative test results). At the end of the visualised period, the middle of August, the government declared that the number of new cases per day was still increasing compared to two weeks ago. It was now clear that most infections occurred at home, for example at a birthday party or a dinner with friends. Therefore regional measures (e.g. wearing facemasks) were implemented (especially in big cities), and people were advised to limit the number of people at activities (e.g. birthday parties) at home to at most six people.

Around 15 August, the prime minister announced the need for people to keep following advices strictly, since more and more people were somewhat neglecting them. A few days later some regional measures were taken in big cities. Together these seemed to make people aware that they really have to take measures seriously for approximately two weeks, since the small increase that started before 15 August stopped for a while. On 1 September, when the number of confirmed cases per started to increase again, few national measures were implemented to prevent a second peak from happening. These included prohibiting people to sing or scream when in groups and starting to educate personal in caring houses about COVID-19. On 20 September, additional regional measures were implemented. People were not allowed to meet with more than 50 people outside anymore and catering services had to close at 1.00 AM. One week later, more measures were implemented that were active for at least three weeks. This included reducing the maximum number of people at home and outside, closing catering services at 22.00 PM, sport games only allowed without audience, and recommendation of wearing facemasks in risk areas (big cities). Unfortunately these measures were not enough to prevent spread. At the beginning of September the number of confirmed cases per day increased exponentially until the middle of October. The number of confirmed cases becomes around six times as high than in the first peak.

Around the highest point of the second peak, on 13 October, strict measures were added to current measures. Facemasks were recommended in all public areas. Catering services were closed, all events and drinking alcohol became prohibited, and performing sports was only allowed with a maximum of four people. After these new measures, the number of confirmed cases per day decreased. The level of confirmed cases before the start of the second peak was however not reached. The number of confirmed cases per day fluctuated around 5,000 for halve a month, and increased again at the beginning of December.

Third peak

The third peak that followed soon after the second peak reached a higher number of confirmed cases per day than the second peak. This peak lead the government to implement a lockdown measures on 15 December. The effects of the second lockdown seem to be less effective than the effects of the first lockdown. This might be caused by the fact that catering services were closed and the maximum number of people at gatherings were small for quite some time already. Or a reason could be that people were getting more and more tired of all restrictions caused by the virus.

One month after the third peak, around the middle of January, the number of confirmed cases per day still fluctuates around 6,000 per day. With a new and more contagious variant of COVID-19 entering at the end of December, spread of the virus remains worrying (RIVM, 2020m). Fortunately, vaccination in the Netherlands started on 8 January which provides hope for the pandemic to finally come to an end.

2.1.2 Development of infected cases: Asymptomatic

What is important to keep in mind is that in the early period of the pandemic, testing capacity was limited. This means that infected cases with mild symptoms or no symptoms at all were not tested in this period. The proportion of asymptomatic infections determined in this period is therefore likely to be biased. However, it is important to identify the number of asymptomatic infections since these cases might be hidden drivers of the virus (Nikolai, Meyer, Kremsner, & Velavan, 2020).

The definition of an asymptomatic case that is used by the WHO is an infected person without overt symptoms and who has been laboratory-confirmed COVID-19 carrier. Risk of spread of the virus increases when the size of the asymptomatic population becomes bigger, which means that it is very important to make people aware of preventive actions like washing hands and restricting traveling (Peirlinck et al., 2020). To be able to minimize risk of the spread, early differentiation between pre-symptomatic and symptomatic infection is important to discover the true proportion of symptomatic versus asymptomatic cases.

Estimating the proportion of asymptomatic cases

Until now, the actual proportion of asymptomatic cases with COVID-19 is still uncertain. Recent evidence is suggesting that early identification of a pre-symptomatic case can be found in elevated serum/plasma lactate dehydrogenase levels, which may facilitate early differentiation. Yet with current data it is still hard to make a proper estimation. Early estimations are said to be between 18% and 81%, which is a very broad range. Data of a comparison study is showing that characteristics of pre-symptomatic and asymptomatic infections are not the same and that younger aged people are more often showing asymptomatic or mild infections, which suggests them to be the (main) hidden spreaders of the virus. However, since the average age of COVID-19 infections is far above the age of children, their role in transmission is not clear yet (Nikolai et al., 2020).

Recent antibody prevalence studies show an increasing amount of evidence for the number of unreported asymptomatic cases. The number of asymptomatic cases could even outnumber the confirmed symptomatic number of cases by an order of magnitude. This was observed in a study that was conducted in New York, where the number of confirmed cases was 10 times less than the number of cases with antibodies (Worldometers, 2020b). This could have been the reason for preventive strategies to fail. One study performed at the cruise ship Diamond Princess provides a proper estimation of the asymptomatic proportion due to the number of people and the tracking of asymptomatic cases. The outbreak of COVID-19 on the cruise ship led to 712 of 3711 persons being infected with the virus. At the time of testing, 58% of these infected persons were identified as being asymptomatic. The majority of those people remained asymptomatic (Sakurai et al., 2020). In the remainder of this report, we assume the estimate of Katri Manninen to be the best estimate for the proportion of asymptomatic cases (see Figure 1). This means 40% of all infected cases are assumed to be asymptomatic. Besides, we assume the development of the number of asymptomatic cases to be similar to the development of symptomatic cases.

2.1.3 Development of the number of hospitalized patients

One of the main reasons of the Dutch government to implement lockdown measures is to prevent hospitals from overflowing. Therefore the number of hospitalized patients is considered as an important indicator for the spread of COVID-19. In Figure 5, the number of hospitalized patients per day can be found. The graph was gathered from the RIVM (RIVM, 2020f).

We observe that peaks occurred around the same time as the peaks that were observed for the number of confirmed cases per day. What is remarkable is that the first peak reaches a point that is almost twice as high as the highest point of the second peak and the third peak. This might indicate that the number of infected cases during the first peak was actually higher than in the second and third peak. However, there may also be other reasons for this observation (for example a decline of the hospitalization rate of infected cases).

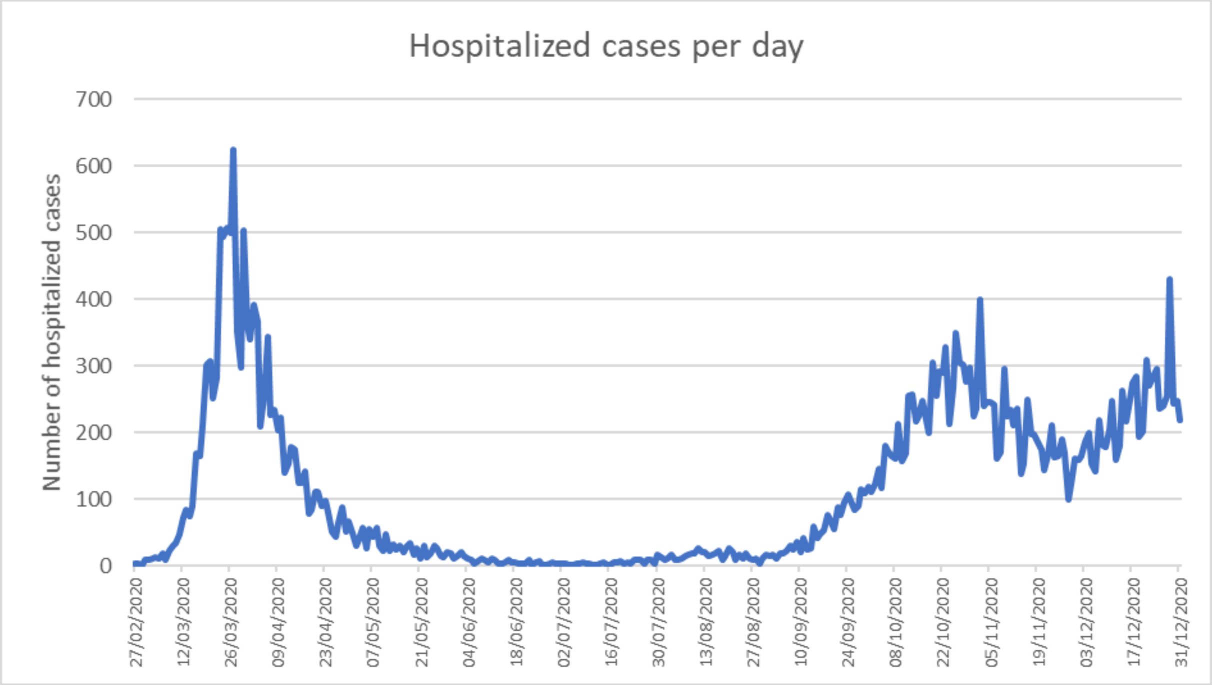


Figure 5 Number of hospitalized patients per day in the Netherlands

2.1.4 Development of the number of occupied IC beds

To be able to get insight in the experienced pressure at the Intensive Care of hospitals, the number of occupied IC bed admissions per day is provided in Figure 6. This number includes patients that were hospitalized in German hospitals when Dutch IC's experienced overflow. The development of the number of IC bed admissions per day is very similar to the number of hospital admissions per day (NICE, 2020).

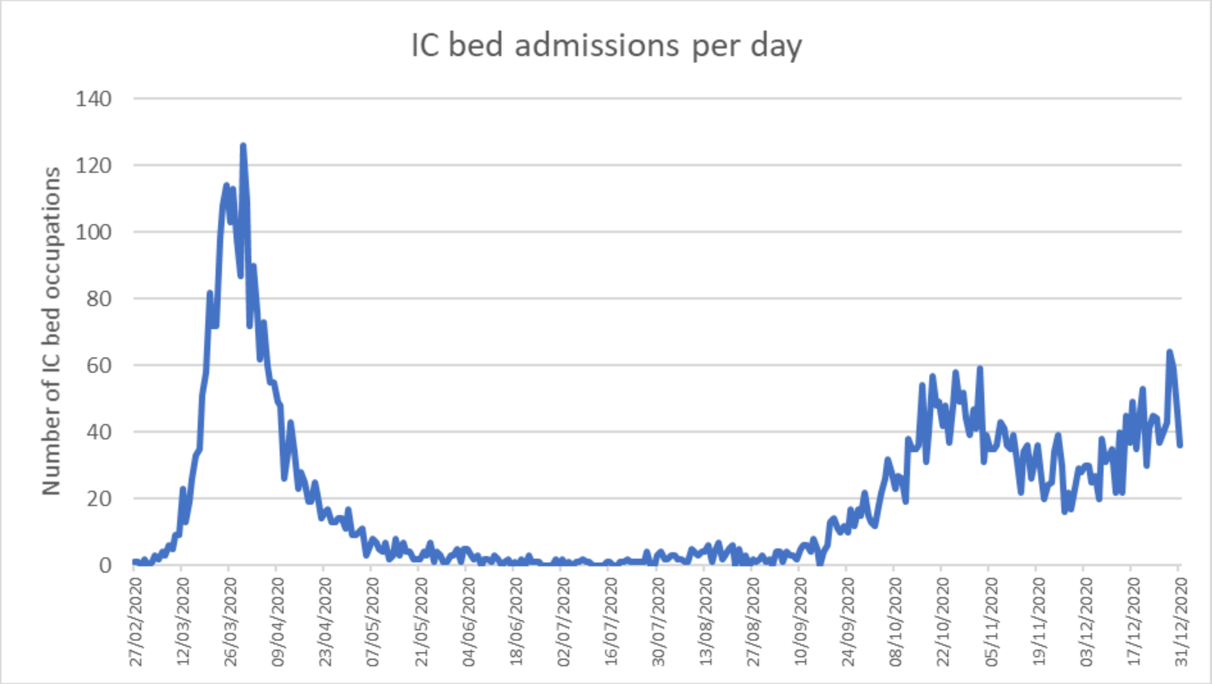


Figure 6 Number of IC bed admissions per day in the Netherlands

2.1.5 Development of the number of deaths

Around one week after the exponential increase of number of confirmed cases, the number of deceased people in the Netherlands started to grow exponentially (see Figure 7). The number of deaths per day, similar to the number of hospitalized patients and the number of IC bed admissions, did not increase together with the number of confirmed cases in the middle of July. We do observe a second and third peak. These peaks are more than half the first peak. In Section 2.2, we discuss possible reasons for this difference.

Figure 7 Number of reported deaths per day in the Netherlands

2.2 Differences between the development of indicators

As we could see in Section 2.1, all indicators show a rather similar development. The biggest difference is that the number of confirmed cases in the first peak is much lower than in the second and third peak, while for the other indicators we see the opposite. It is important to find out the reason for this, since this may help to discover how we can prevent hospitals from overflowing. We consider four possible reasons for this difference in development of indicators between the peaks. Firstly changes in age-distribution, which means that the distribution of the age of infected persons in the first peak differs from the age-distribution of infected persons in the second peak. Secondly changes in gender-distribution. Thirdly changes in hospital pressure, since there was more pressure in hospitals during the first peak than in the rest of the period. And lastly, changes in the number of tests. Another reason for the difference in development of the number of confirmed cases and the other indicators could be that the virus transformed between the two peaks. Since a considerable transformation of the virus only entered the Netherlands at the end of December, we do not include this possibility in our scope.

Reason 1: Changes in age-distribution

In Figures 8, 9 and 10, the age- (and gender-) distribution of confirmed cases, hospitalized patients and deaths in the Netherlands are visualized respectively. These figures originate from the weekly COVID-19 update of the RIVM (RIVM, 2020j). In Figure 9, the distribution of the number of hospitalized patients shows that older people are more likely to be hospitalized. In the distribution of deaths (Figure 10) this relation is even more obvious.

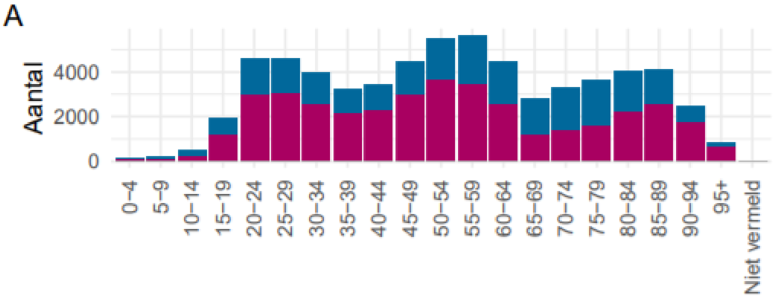


Figure 8 Age-gender distribution of confirmed cases in the Netherlands (from 27 February to 17 August)

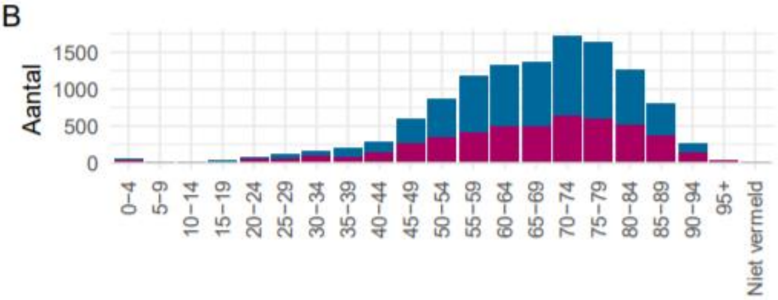


Figure 9 Age-gender distribution of hospitalized cases in the Netherlands (from 27 February to 17 August)

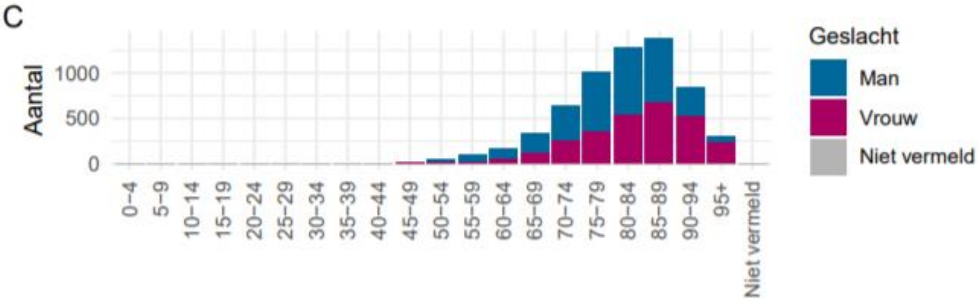


Figure 10 Age-gender distribution of deaths in the Netherlands (from 27 February to 17 August); “Aantal” = Number, “Geslacht” = Gender “Man” = Male, “Vrouw” = Female, “Niet vermeld” = Not identified

In the previous section, we observe that the first peak of the number of confirmed cases is much lower than the second and third peak, while for the other indicators the first peak is once as high. Based on the findings above we conclude that when the average age of infected cases decreased between the first and second peak, this results in a lower number of hospitalizations, IC occupations and deaths with the same number of confirmed cases. One study from the ECDC including infected cases from all over Europe indeed confirms that during the first peaks the age of infected persons was higher than a few months later. In Europe, 40% of the cases were aged 60 years or older and most cases were between 50 and 59 years old between January and May. In the months June and July, 17.3% of the cases were aged 60 years or older and most of the cases were aged between 20 and 29 years (19.5%) (ECDC, 2020). The effect of changes in age-distribution on spread will be studied in more depth in the literature review (Chapter 3).

Reason 2: Changes in gender-distribution

Next to the distribution of age, the distribution of gender is visualised in Figures 8, 9 and 10. We see that women are more likely to be infected (or are tested more often), but that men are more likely to be hospitalized or to die. Since we did not find any studies about changes in gender-distribution between February and December, we leave the effect of gender on spread out of our scope.

Reason 3: Changes in hospital pressure

During the first peak, hospitals were struggling with their available capacity. This could have caused differences between the development of indicators, because hospitals were having more sufficient capacity during the rest of the period. Hospital capacity can especially influence the number of deaths, because the number of deaths follows from provided healthcare. The effect of hospital pressure will be studied in more depth in the literature review.

Reason 4: Changes in number of tests

The number of tests performed per day significantly increased between 27 February and 31 December. Where the number of tests performed per week in the first week of June was around 40,000, at the end of December this number had increased to above 400,000 per week. At the beginning only people with chance of a severe course of the disease and hospital personnel were able to be tested. This contributed in significant differences in the number of confirmed cases between the first peak and the second and third peak. The effect of the number of tests on spread will be studied in more depth in the literature review.

2.3 Differences between provinces

A drawback of existing models that aim to predict the spread of COVID-19 is that often average quantities are used to model spread in a large area. However, individual-specific control measures show to outperform population-wide measures in an outbreak (Lloyd-Smith, Schreiber, Kopp, & Getz, 2005). This means that it can be particularly important to predict the spread on local level instead of national, such that estimations of parameters can be more precise. We scale down the spread in the Netherlands to spread in provinces to make sure that enough data is available. Unfortunately, there is far less data available from provinces in specific.

The number of confirmed cases, the number of hospitalizations and the number of deaths reported per municipality are visualized in Appendix A. These graphs are gathered from the RIVM (RIVM, 2020a). We took three different moments to explain the differences in spread between provinces over time. These are the start of the first peak, the end of the first peak, and the start of the second peak in number of confirmed cases.

The differences of spread between provinces in the Netherlands changed a lot over time. On 10 March, the RIVM posted a news report about the differences in infections between provinces. At that moment, there were 23,097 confirmed cases in the Netherlands, of which almost a quarter living in the province Noord-Brabant. The number of hospitalizations and deaths was also highest in Noord-Brabant. Thereafter the highest number of confirmed cases, hospitalizations and deaths were observed in Limburg, and in a small part of Gelderland and Zeeland (see Appendix A). Groningen showed the least number of infections (1.1%), followed by Drenthe (1.2%), Friesland (1.4%) and Flevoland (1.4%). The number of confirmed cases in Noord-Brabant seemed to flatten after a while and the number of confirmed cases in the northern part of the Netherlands remained low (RIVM, 2020k). At the end of the first peak, the spread was far more equally spread over the entire country. Only the northern part of the Netherlands showed a relatively low number of confirmed cases, hospitalizations and deaths.

In Section 2.1.1 we already mention that the start of the second peak was mostly caused by infections in bigger cities, different from the spread during the first peak. The provinces Noord-Holland, Zuid-Holland, and Utrecht (and one municipality in Friesland) were affected most at the beginning of the second peak, which is where bigger cities are located. Fortunately, the number of hospitalizations and deaths in the whole country remained low during this time. This indicates that the Netherlands could gain profit by implementing local control measures to prevent further spread. By what extent this is true in practice could be discovered by applying different measures in different provinces in the system dynamics model. However, during the second and the third peak the number of cases in the rest of the Netherlands started to increase as well. At the end of the third peak, the virus had spread everywhere in the Netherlands. For this reason, we do not test the effect of applying regional measures in this study.

2.4 Conclusion

Now we answer research question 1, “how did the spread of COVID-19 develop in the Netherlands?”.

Between the first infected case on 27 February and 31 December, the number confirmed cases per day showed three peaks. The first peak was flattened by implementing an intelligent lockdown. After a decline in number of cases due to this intelligent lockdown, measures were gradually softened. The number decreased to below 200 new cases per day and remained there for quite some time. The number of confirmed cases per day started to increase again around 1 July, when bigger events and sports were allowed again to a certain extent. At first sight, this increase in confirmed cases did not lead to higher values of other indicators. But after some time, the observed number of confirmed cases reached 10,000 and the number of hospitalizations, IC occupations, and deaths per day increased as well. While the number of confirmed cases in the second peak and the third peak were much higher than in the first peak, the number of hospitalizations, IC occupations, and the number of deaths observed during the second and third peak were approximately half of the observed values during the first peak. In the middle of December, lockdown measures were implemented for the second time to prevent the virus from spreading.

Four possible reasons for the difference in peaks of indicators were discussed in Section 2.2. We observe that a change in age-distribution and a change in number of tests per day can significantly influence the development of indicators. This is assumed because there is evidence for these two factors to have changed within the studied period. In the second peak of confirmed cases, the age of infected cases is assumed to be lower than the age during the first peak. The number of tests performed per day increased with time, causing more and more people with mild or asymptomatic symptoms to be able to be tested. The effect of both age and number of tests on spread will be studied in more depth in the literature review. A change in hospital pressure can also be a reason for the difference in peaks of indicators, since hospital pressure during the first peak was higher than the experienced hospital pressure thereafter. Hospital pressure depends on the available and utilized hospital capacity. We identify in the literature review whether hospital pressure considerable affects spread. The gender of an infected case seems to influence the course of the disease, but we did not find evidence of a change in gender-distribution between February and December. Therefore this last reason is not taken into account in the scope of this study.

Besides differences in spread between indicators, we observe differences in spread within the Netherlands. During the first peak, provinces Brabant and Limburg were mostly affected. The second peak started in bigger cities in Noord-Holland, Zuid-Holland, and Utrecht. During the second and third peak, the virus had spread over the whole country. Implementing regional measures is therefore not considered to be an effective strategy at this moment in time.

3. Literature research

In this study, a system dynamics model will be built to be able to discover effective measures for preventing spread of COVID-19 in the Netherlands. Factors and measures that have substantial impact need to be identified to develop a reliable model. Key factors and measures will be used as input parameters for the model. In this chapter, we first define the factors we are going to study. Relevance of these factors will be determined with help of a literature review in Section 3.2. In Section 3.3, we identify and explain the measures that are considered to prevent spread of the pandemic. And in Section 3.4, we determine the approach that we use to build the system dynamics model.

3.1 Factors influencing the spread of COVID-19

As mentioned in Chapter 1, the WHO expressed the urgency to understand the relative importance of different routes of transmission required to develop infection prevention measures. Besides, the WHO suggests that the virus spreads primarily via contact routes and respiratory droplets. The WHO finds it most important that research is performed to among others risk factors, settings of superspreading events, and the extent of transmission both pre-symptomatically and asymptomatically. One major reason that much is still unknown about transmission is caused by the highly variable characteristics and behaviour of viruses, because each virus acts different (Wigginton & Boehm, 2020). This means that it is hard to learn from other viruses how SARS-CoV-2 spreads through the environment.

For the reason above and the availability of data, the following factors are studied for their importance in the spread of COVID-19. We divide the factors in three types: Disease, demographics and other.

Disease:

- Incubation period
- Infectious period
- Reproduction rate
- (Case-) Fatality ratio

Demographics:

- Age

Other:

- Weather (temperature, humidity, wind speed)
- Contact rate (population density, adoption of government measures, places of infection)
- Testing capacity

All parameters will be estimated with help of data from the period between 27 February (the first confirmed case) and 31 December. When considered necessary, we provide differences in estimates between provinces. For some factors, there may exist some delay that has to be included in the model. For example, delay due to incubation time or testing time. The literature review will help to discover possible delay and to discover how delay can be included in the model.

3.2 Relevance of the factors

In this section, values of factors and their relevance will be determined with a literature study. As we mention in Section 2.2, values of indicators are changing differently over time. For this reason, estimates of some factors may deviate for different moments in time. For these factors we provide necessary information to give insight in these deviations.

Factor 1: Incubation period

The incubation period is the time between the moment of infection and the moment symptoms are showing. The incubation period of COVID-19 is estimated to be on average five to six days, but can be between 2-14 days (RIVM, 2020d). We assume an average incubation period of 5.5 days in our model. The testing and reporting delay for infections is between 1 and 4 days, for which we take an average of 2 days. Incubation time, testing delay and reporting delay together cause a person to be infected on average seven days before the infected case becomes a confirmed case.

Factor 2: Infectious period

The start of the incubation period is different from the start of the infectious period. The infectious period is the time period in which an infected person can infect others. There is not (yet) enough information about when an infected person is in his or her infectious period. In general it is assumed that an infected person can infect others when the infected person is showing symptoms, but there are indications for transmission of SARS-CoV-2 before symptom onset. The role of asymptomatic cases in spread of the virus is still not fully understood (RIVM, 2020d).

Viral detection

Several studies show that 1 to 3 days before an infected person is showing symptoms, the virus can be detected. The amount of virus proven in the infected person is highest at the moment symptoms begin to show and gradually declines thereafter. However, detection of the virus does not necessarily mean that the person is able to transmit the virus. Viral detectability could cause overestimation of the infectious period between two and six days. The time period that the virus can be detected differs based on the severity of the disease. For asymptomatic cases, this turns out to be approximately 1 to 2 weeks, for patients with mild to moderate disease for up to three weeks and for patients with severe disease this can be even longer (WHO, 2020c). Some studies reported an association between severity of disease and duration of the infectious period, though evidence for this relation is not strong (Byrne et al., 2020).

Duration of infectious period

Byrne et al. (2020), who study the infectious period with a literature review, suggest the duration of the infectious period of asymptomatic infections to be between 4 and 9.5 days. This distribution could, due to viral dynamics, have a longer tail with low probability for up to 20 days. In another simulation study, the infectious period for asymptomatic cases is assumed to be 5.76 days (Peirlinck et al., 2020). The infectious period of symptomatic cases is in several papers estimated to start approximately two days before symptom onset with a range between one and five days (RIVM, 2020d). The proportion of transmissions before symptom onset was estimated to be 44%, yet the accuracy of this estimate has to be questioned due to a lack of data. The infectious period continues for up to seven days from the onset of symptoms with a peak at 0.7 days (Byrne et al., 2020).

Since the precise duration of the infectious period is still unknown, we estimate the duration of the infectious period. From Chapter 1, we find that a mild form of the disease is shown in about 85% of the cases, of which we assume 40% to be asymptomatic. Serious symptoms are shown in approximately 10% of the cases and a critical condition develops in around 5% of the cases. For asymptomatic cases the infectious period is assumed to be 6 days, for moderate symptomatic cases 9 days, and for severe symptomatic cases 15 days. This gives 40% of asymptomatic cases who have an infectious duration of 6 days and 60% of symptomatic cases for whom we assume an infectious period of 11 days. This leads to an average infectious period of 9 days. In the system dynamics model, the infectious period of all cases is assumed to be the same. The duration of the infectious period will only be distinguished based on severity of disease when we consider both symptomatic and asymptomatic cases, which we do in the sensitivity analysis.

Factor 3: Reproduction rate

The reproduction rate is a way to indicate spread of a virus, and thus can help to define the outbreak dynamics. The reproduction rate R shows the rate by which the virus spreads. It indicates how many people are infected by one infected person on average. With a rate of 1, the number of infections per day will not change over time. When the rate is lower than 1, the number of infected persons will decrease and when the rate is higher than 1 the number of infections will increase. The effective reproduction rate is implemented in our model as the rate of transmission.

Effective reproduction rate

The extent by which an infectious individual transmits the virus is indicated with the reproduction rate. Yet there is a difference between the basic reproduction rate (R_0) and the effective reproduction rate (R_e). The first applies when no outbreak control is applied to limit spread, and defines the rate as the average number of cases produced by an infected individual in a fully susceptible population (He, Yi, & Zhu, 2020). R_0 is initially assumed to be 2.5 in our model. This is in accordance with information from the RIVM, who estimate R_0 to be between 2 and 2.5 (RIVM, 2020c). The basic reproduction rate will eventually be calibrated in Chapter 4. The effective reproduction rate is influenced by external factors like contact rate and government measures (Peirlinck et al., 2020).

Development of the effective reproduction rate in the Netherlands

In first instance, the reproduction rate is calculated with help of the number of hospitalized patients. When this number is very low, there is not enough data to calculate a reliable reproduction rate and the number of confirmed cases is used instead. In Figure 11 and 12, the reproduction rate for the Netherlands is visualized (indicated with the dark purple line). These are estimations of the RIVM (RIVM, 2020j). The rate is determined based on hospitalizations on the left side of the dotted line (before 15 June). On the right side of the dotted line the rate is based on the number of infections

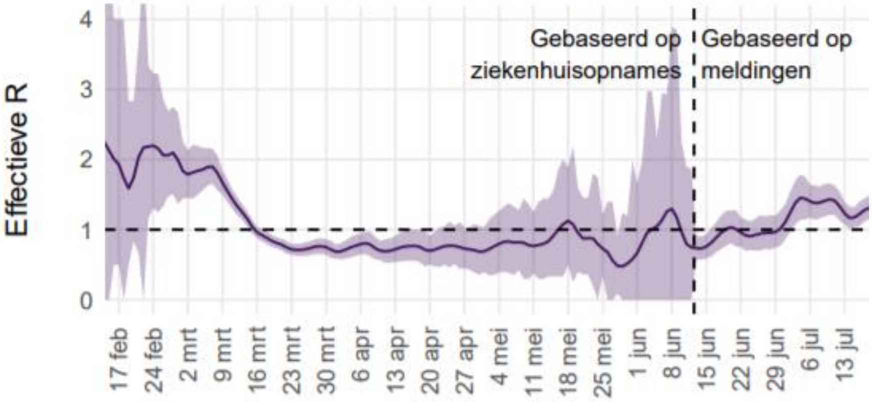


Figure 11 Development of the reproduction rate in the Netherlands from 17 February until 13 July

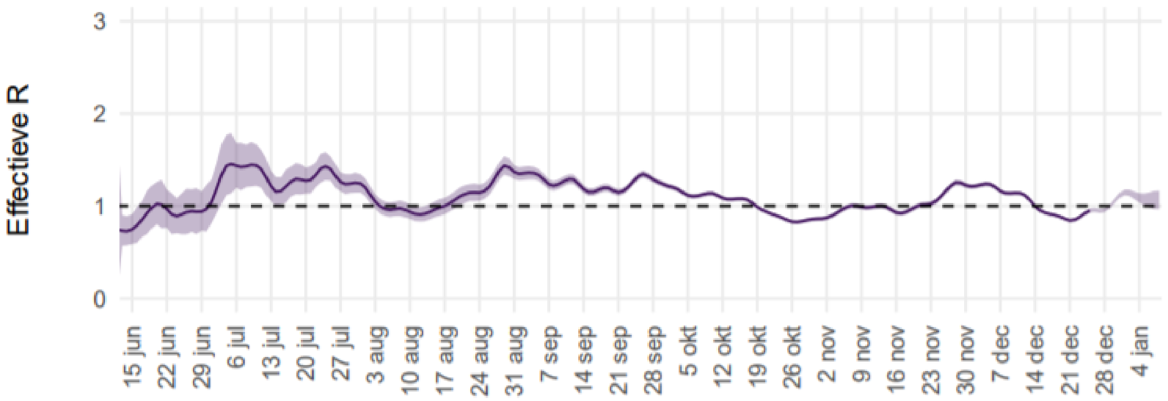


Figure 12 Development of the reproduction rate in the Netherlands from 15 June until 28 December

(after 15 June). The light purple area indicates the bandwidth of the rate. When there is a limited amount of data available, the uncertainty of the rate is higher and the bandwidth of the real rate will become wider. Estimations of the rate are most accurate with data from 14 days ago, due to testing and reporting delay and time between successive infections. For this reason, the rate (dark purple line) in Figure 11 stopped around 23 December.

As we can see in Figure 11, the reproduction rate was around 2 when COVID-19 started to spread in the Netherlands with a confidence interval from 0 to around 4. The rate remained above 1 until the middle of March. When comparing Figure 2 to Figure 11, we observe that a rate below 1 turned the increasing number of confirmed cases per day into a horizontal line. From the middle of March until the end of June the rate remained mainly below 1. At the beginning of July the rate increased above 1 again. This caused the number of confirmed cases per day to increase again around the middle of July. From the beginning of August to the middle of August the rate declined to 1 and increased again after that. We learn from this is that the reproduction rate is an effective indicator for the prediction of the spread of the virus.

Over- and underestimating the rate of transmission

Just like for many other infectious diseases, the true level of transmission for COVID-19 is often underestimated. This is due to people who have been infected but are not detected as infected and thus not counted as infected. Effects of underestimation may even be enhanced during an epidemic, when testing capacity is limited. When dealing with limiting testing capacity, people in risk groups are more likely to be tested (WHO, 2020a). The effects of testing capacity will be taken into account with the factor “testing capacity”.

Effective reproduction rate in the system dynamics model

To be able to determine the number of infected cases in our system dynamics model, we have to predict the effective reproduction rate. Since we want to determine the effect of relevant factors and measures with our model, we will develop a regression model that predicts the effective reproduction rate based on the values of relevant factors and measures. This model is built in Section 4.3.

Factor 4: Fatality ratio

The fatality ratio is an important characteristic of an infectious disease. The ratio helps to understand the severity of the disease, can evaluate the quality of healthcare and identifies populations at risk. There exist two types of fatality ratios, the *infection fatality ratio* and the *case fatality ratio*. The infection fatality ratio (IFR) estimates the proportion of deaths among all infections, while the case fatality ratio (CFR) estimates the proportion of deaths among confirmed cases. The IFR is more accurate, but harder to determine.

Calculating the case fatality ratio

Since there exists no reliable data for the number of infected cases, we only calculate the CFR. We use the calculated CFR as indicator for the fatality ratio in the Netherlands. The CFR can be calculated with the following formula:

$$Case\ fatality\ ratio\ (CFR, in\ \%) = \frac{Number\ of\ deaths\ from\ disease}{Number\ of\ confirmed\ cases\ of\ disease}\ x\ 100$$

At present, it is very likely that a "misleading" estimate of the CFR will be made. A precise CFR can best be determined when the epidemic has ended (Worldometers, 2020b). Anderson, Heesterbeek, Klinkenberg & Hollingsworth (2020) suggest that the best estimate for the IFR of COVID-19 should be between 0.3% and 1% (Anderson, Heesterbeek, Klinkenberg, & Hollingsworth, 2020). When all infected cases would be confirmed cases, the CFR would be equal to the IFR.

We calculate the CFR for each day between 1 April and 17 November. After 17 November, we do not have sufficient data available yet. In our calculation, we take the number of deaths of today and the number of confirmed cases of 3 weeks ago to determine the CFR of today. With these three weeks we take into account reporting delay (assumed 2 days), incubation time (2 – 14 days), and duration of the disease before death. These three weeks lay inside the range of Katri Maninnen (see Figure 1).

Development of the case fatality ratio in the Netherlands

In Appendix B, the development of the estimated CFR is visualised in a graph. For this calculation we use the number of deaths reported to the RIVM. However, the number of deaths reported to the RIVM is not the real number of deaths caused by COVID-19. According to the CBS (Central Agency of Statistics), the number of reported deaths to the RIVM from March to June (6,115) was 3,952 less than the number of deaths reported to the CBS (10,067), which is a considerable difference (CBS, 2020). We do not take this difference into account in our model.

We observe the CFR fluctuates a lot and see a decreasing trend from approximately 10% to 1% between April and November. According to an article of NU.nl, fluctuations in the ratio are caused by a reporting delay causing one day with many reported deaths (NU.nl, 2020f). In reality we assume that deaths are spread more equally over the week. Due to multiple factors that can cause over- or underestimation, the ratio might be biased at times. This is the reason for the large difference in CFR we observe in Appendix B. Below we shortly discuss considerable factors for over- or underestimation.

Over- and underestimating the case fatality ratio

It is important to notice that a fatality ratio is sensitive for over- and underestimation. Bad timing of the calculation can be a reason for this. For example, overestimation of the CFR was particularly the case at the beginning of the pandemic, when testing capacity was low and milder or asymptomatic cases were less likely to be identified as a confirmed case of the disease. Besides testing capacity, patient characteristics and hospital capacity show to influence the ratio and are thus important to take into account.

Age groups

The higher the age of a person, the higher the CFR turns out to be (Wu & McGoogan, 2020). The case fatality ratio of COVID-19 was estimated to be 2.3% by Wu & McGoogan, where people between 70 and 79 years show a ratio of 8%, people older than 80 years a ratio of 14.8%. While this is an early approximation of the CFR and likely to be biased, these differences between age-groups provide an insight in the effect of age on the CFR.

Hospital capacity

The number of deaths was particularly high during the first peak when hospitals were struggling with their available capacity. Therefore we consider pressure at hospitals to be able to influence the number of deaths. Normal capacity of the IC in the Netherlands is 1,150 beds (Venticare, 2020). Capacity can be extended when the need is high, which was the case during the first peak of infections. This way hospitals were able to take care of all patients and the fatality ratio is not assumed to have increased between February and December in our model.

Testing capacity

Since before 1 June not every Dutch resident was allowed to be tested, we consider the CFR before this date to be biased. After 1 June, the number of tests per day gradually increased until December. Due to utilization of testing capacity we consider the CFR in November to be the most accurate CFR for this study.

Fatality ratio in the system dynamics model

Since the number of tests per day gradually increased between April and November, we assume the average CFR in November to provide the most accurate estimate for the IFR. The average CFR of November is 0.9%, which lays inside the suggestion of Anderson et al. (2020). We will use the IFR to determine the number of deaths in the system dynamics model. The CFR will be used to express the fatality ratio, because this ratio is easier to measure since there exists publicly available data about the number of confirmed cases (WHO, 2020a). The correctness of our ratio will be investigated by calibrating the system dynamics model.

Factor 5: Age

In Section 2.2 we conclude that the age of an infected case influences the rate of a case to be hospitalized or to die. Below we provide the age-distribution for confirmed cases, hospitalizations, and deaths, and the differences in age-distributions between provinces, to determine the relevance of this factor on spread. The age-distribution of indicators will be determined for three intervals to provide insight. For the number of hospitalizations and the number of deaths these intervals are the start of the first peak (end of March), the period between first and second peak (end of April), and the start of the second peak of infections (beginning of August). For the number of confirmed cases we determine the age-distribution for three moments after 1 July, because there is no publicly available data from before July. Next to the three above indicators, we find the age-distribution of patients who occupy an IC bed relevant, which we assume to be the same as the age-distribution of hospitalized cases.

Distribution of the number of infected cases over age groups

Unfortunately, there is a limited amount of publicly available data about the age of infected cases. For this reason, we observe historical data of the number of confirmed cases from the RIVM of three moments, with approximately one month in between. Data is separated per age-group and can be found in Table 1. We assume that mostly higher aged individuals were infected before July, with a gradual change to lower aged individuals. The highest number of infections on 24 March is seen in the age-group 55 to 59 years and after that between 70 and 85 years (Ranzijn, 2020). According to historical data of the RIVM, the highest number of infections until 17 August occurred at individuals aged between 50 and 59 (see Table 1). One month later the age-distribution changed with a highest number of infections for individuals aged between 20 and 29. Consequently, confirmed cases at the beginning of the studied period were on average older than at the end, with an average age of 58 years measured on 14 July and 50 years measured on 18 September. We assume this to mostly be caused by the fact that most younger people were not allowed to be tested at the beginning of the studied period, because they were not considered to be part of the risk group. To find out whether this is the case, we study test-adjusted incidence over age groups below.

Table 1 Age-distribution of confirmed cases in the Netherlands

Age-group	**14-Jul**	**17-Aug**	**18-Sep**
0-9	1.0%	1.0%	1.0%
10-19	2.0%	4.0%	6.0%
20-29	10.0%	14.0%	19.0%
30-39	10.0%	11.0%	13.0%
40-49	12.0%	12.0%	12.0%
50-59	18.0%	17.0%	16.0%
60-69	12.0%	11.0%	10.0%
70-79	13.0%	11.0%	9.0%
80-89	16.0%	13.0%	10.0%
90>	6.0%	5.0%	4.0%

Distribution of test-adjusted incidence over age groups

A recent study observed that relative patterns of age, regardless of the epidemic phase, remain identical when adjusting for temporal trends of testing and disease incidence (Fisman et al., 2020). According to Fisman et al. (2020), individuals who are aged older are more frequently tested than younger people due to a difference in severity of the disease. This study compares observed infection incidence with test-adjusted incidence, which is found to considerably influence incidence per age-group (see Figure 13). This means that when we take away effects of testing, the change in age-distribution we observe in Table 1 would probably not have occurred. For this reason, we assume the age-distribution of infected cases did not significantly change between February and December 2020, and observed differences in age-distribution of confirmed cases within the studied period are mostly caused by differences in testing. Consequently, we do not distinguish in age-groups between infected cases in our model.

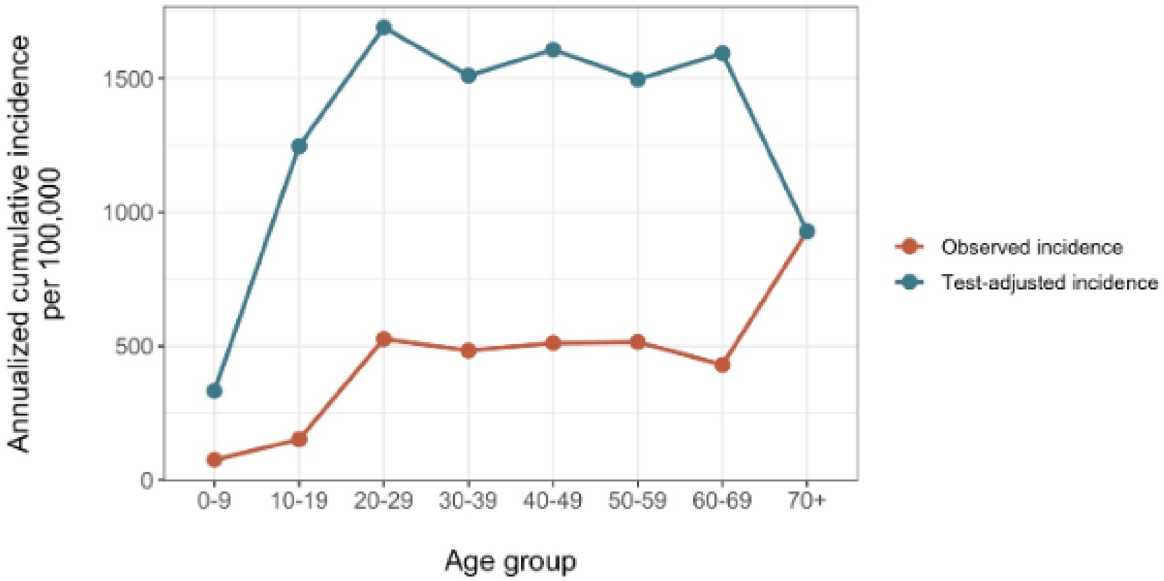


Figure 13 Observed incidence versus test-adjusted incidence per age-group

Distribution of the number of deaths over age groups

We derive an age-distribution for deaths (Figure 14), based on data of the RIVM. In this figure we observe that the number of deaths in the Netherlands is increasing with age with an average age of 81 years, and that the average age of deceased cases increases between the first peak and the second peak. The percentage of cases aged above 80 years increased and the percentage of cases aged below 80 years decreased. We believe main causes for to be a change in reported number of COVID-19 deaths to the RIVM and actual number of COVID-19 deaths between first and second peak, and hospital personnel knowing better how to deal with the disease. While we observe a small change in age-distribution between age groups of deaths, the average age of deaths in the Netherlands remained approximately 81 years between July and September. For this reason, we do not distinguish in age-groups between deaths in our model.

Compared to other countries, the Netherlands shows a relatively low number of deaths under hospitalizations. This is assumed to be due to the level of preparedness in the Netherlands (Aleva, van Mourik, Broeders, Paling, & de Jager, 2020). According to Worldometer, a reference website that provides counters and real-time statistics for diverse topics, the average age of deaths due to COVID-19 worldwide is approximately 79 years (Worldometers, 2020a), which is somewhat lower than observed in the Netherlands.

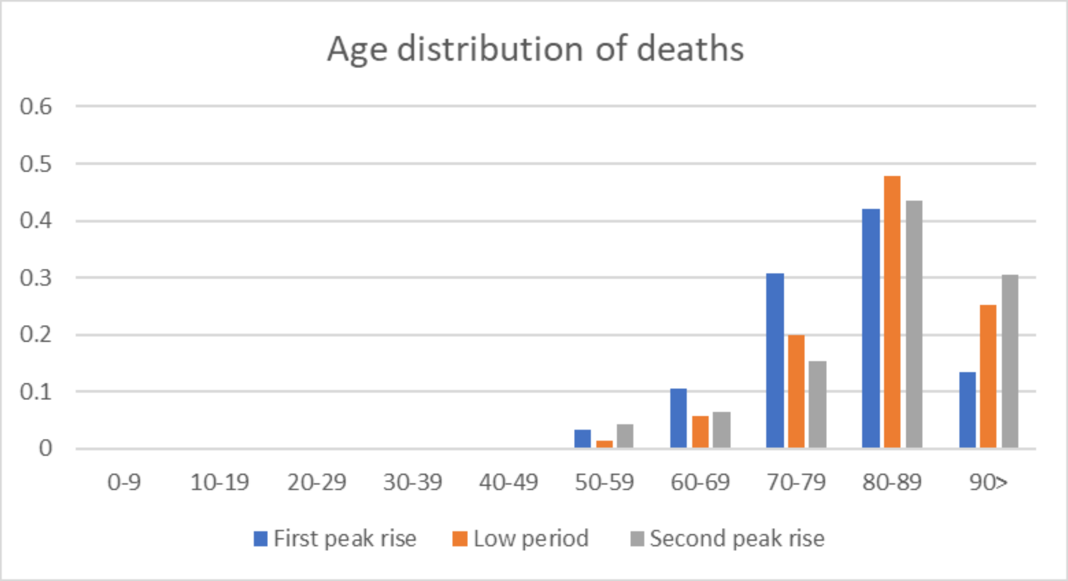


Figure 14 Age-distribution of deaths for different phases

Distribution of the number of hospitalized cases over age groups

We derive an age-distribution of hospitalized cases (Figure 15), based on data of the RIVM. In the histogram we observe a small change in age-distribution between 27 February and 17 August. The percentage of young people of the total number of hospitalizations increases over time, while the percentage of older people decreases a little. However, the average age of hospitalizations in the Netherlands remained 66 years between July and September. For this reason, we assume the age-distribution of hospitalized cases to remain the same within the whole studied period, and we do not distinguish in age-groups between hospitalizations and IC occupations in our model.

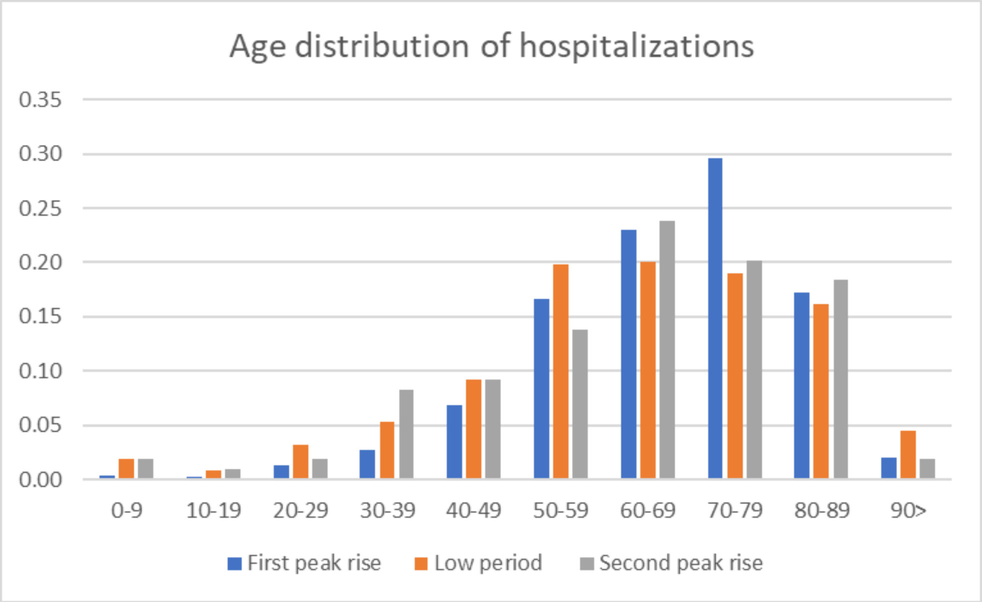


Figure 15 Age-distribution of hospitalized patients for different phases

Differences in the distribution of age between provinces

In Section 2.3, we observe differences in spread of COVID-19 between provinces. During the second peak, the virus mostly spread in bigger cities and the average age of confirmed cases declined. These bigger cities are located in provinces where the percentage of people aged between 20 and 29 is somewhat higher than in other provinces (e.g. Noord-Holland, Utrecht). To be able to take into account differences between provinces that are caused by age, we determine the number of people in a certain age-group (Table 2) and the average age (Table 3), for each province. Data was retrieved from StatLine, an open database of the CBS, and is from May 2020 (Statline, 2020).

Table 2 Percentage of people in age-group per province

	0-9	**10-19**	**20-29**	**30-39**	**40-49**	**50-59**	**60-69**	**70-79**	**80-89**	**90>**
Groningen	0.09	0.11	0.16	0.11	0.12	0.14	0.12	0.09	0.04	0.01
Friesland	0.10	0.12	0.11	0.11	0.12	0.15	0.13	0.10	0.04	0.01
Drenthe	0.09	0.12	0.10	0.10	0.13	0.16	0.14	0.11	0.05	0.01
Overijssel	0.10	0.13	0.12	0.12	0.13	0.14	0.12	0.09	0.04	0.01
Flevoland	0.12	0.13	0.13	0.14	0.13	0.15	0.11	0.06	0.02	0.00
Gelderland	0.10	0.12	0.12	0.11	0.12	0.15	0.13	0.09	0.04	0.01
Utrecht	0.11	0.12	0.14	0.13	0.13	0.14	0.11	0.08	0.03	0.01
Noord-Holland	0.10	0.11	0.13	0.14	0.13	0.14	0.11	0.08	0.04	0.01
Zuid-Holland	0.11	0.11	0.13	0.13	0.13	0.14	0.11	0.08	0.04	0.01
Zeeland	0.10	0.11	0.10	0.11	0.12	0.15	0.14	0.11	0.05	0.01
Noord-Brabant	0.10	0.11	0.12	0.12	0.13	0.15	0.13	0.09	0.04	0.01
Limburg	0.08	0.10	0.12	0.11	0.12	0.16	0.15	0.11	0.05	0.01

Table 3 Average age per province

Province	**Average age**
Groningen	41.80
Friesland	42.86
Drenthe	44.26
Overijssel	41.46
Flevoland	38.49
Gelderland	42.26
Utrecht	39.99
Noord-Holland	41.06
Zuid-Holland	40.82
Zeeland	44.26
Noord-Brabant	42.41
Limburg	45.03

In Table 2 and 3, we observe small differences in age-distribution between provinces. In first instance we will not take differences in age-distribution between provinces into account, since the main focus of this study is to analyse the spread of COVID-19 in the Netherlands.

Factor 6: Weather

Whether the weather plays a role in the transmission of the virus is not yet clear. Since the virus is spreading world-wide, it is assumed that seasonal factors do not significantly influence the transmission of COVID-19. Yet there is evidence that exposure to ozone and periods with higher environmental temperature can worsen heart-, lung-, and airway disorders. If it can also deteriorate the course of COVID-19 disease is not known (Warmedagen, 2020). To discover whether seasonal factors influence the spread of the virus, we study the relation between the spread and the sub-factors temperature, humidity, and wind speed. The development of these factors is outlined below, with help of data from the KNMI (Dutch Royal Meteorological Institute)(KNMI, 2020). We assume the same values everywhere in the Netherlands. The temperature, humidity, and wind speed provided below were measured in the Bilt, a municipality in the middle of the Netherlands.

Temperature

The average and maximum temperature per day can be found in Figure 16. We see in this figure that temperature is steadily increasing towards summer and steadily declines after summer holidays. All peaks of the number of confirmed cases were observed when temperature was relatively low, which might indicate that a lower temperature has a positive effect on the spread. According to a systematic review, low temperature and low humidity are potentiating factors of spread of the virus (Mecenas, da Rosa Moreira Bastos, Rosário Vallinoto, & Normando, 2020).

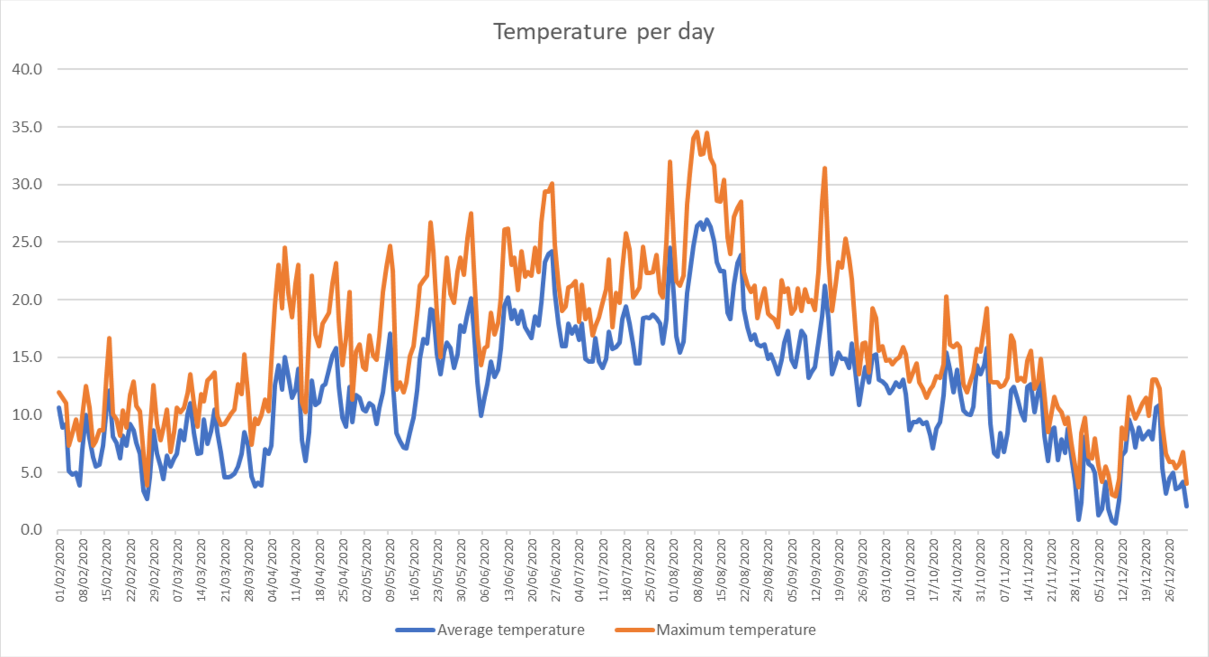


Figure 16 Average and maximum temperature per day in the Netherlands

Humidity

The air can only contain a limited amount of moister. The amount of moister depends on the temperature. Both air containing a lot of moister and air containing very little moister can be perceived as uncomfortable (KNMI, n.d.). The amount of moister in the air is indicated by humidity. The mean relative humidity per day in the Netherlands is visualized in Figure 17.

Air humidity can cause people to perceive the real temperature as another temperature ("sensed temperature"). The higher the humidity, the higher sensed temperature can become at the same real temperature. This can be dangerous for people, since it can lead to overheating.

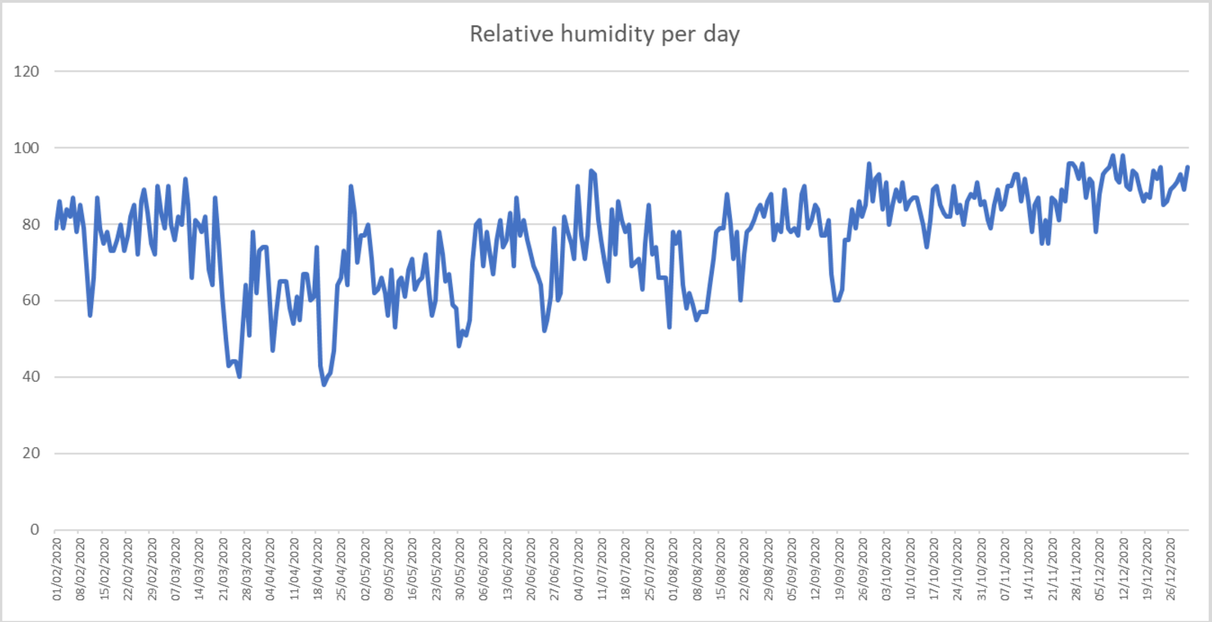


Figure 17 Relative humidity per day in the Netherlands

People have to maintain their body temperature at approximately 37 degrees, but regulation of body temperature becomes harder when the environmental temperature and humidity increase. A temperature of 30 degrees Celsius with a humidity of 50% feels like 31 degrees, but with a humidity of 80% this can feel as 38 degrees. Whether a high sensed temperature leads to an increase in spread will be identified in the statistical analysis. If this is the case, temperature and humidity together provide a better prediction for the spread of the virus than one of the two on its own.

Wind speed

Wind speed is measured by the KNMI in 0.1 m/s. The scale of Beaufort is used to express the speed of the wind. The mean wind speed on a day can be found in Figure 18. Based on the information in the figure it is hard to say something about the relation of wind speed and spread of the virus. With help of statistical analysis we identify whether there exists a relation between spread and wind speed.

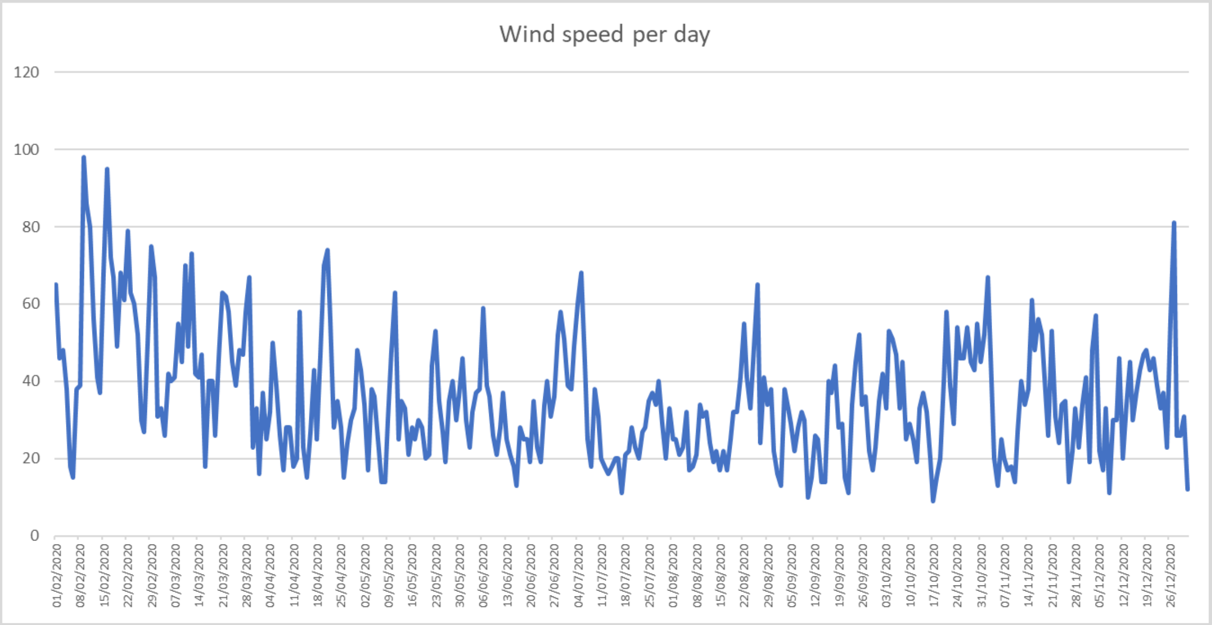


Figure 18 Average wind speed per day in the Netherlands

Factor 7: Contact rate

Contact rate indicates the extent by which a person is in contact with other persons. This rate significantly influences the way by which COVID-19 spreads, since people who have more contact with other people are more likely to infect or be infected. Due to the severe impact of contact rate, we identify ways to express this factor and consequently to be able to include contact rate in our model. We consider a positive correlation between contact rate and three sub-factors: Population density, adoption of government measures, and places of infection. Below, these sub-factors are quantified and their development between February and December is described. These sub-factors will be used to express contact rate in the remainder of this thesis.

Population density

The population density per province can be found in Table 4. The (average) population density in the Netherlands is 517 residents per km². We can see that there are big differences in number of residents between the provinces. When we take the population density of a province relative to the average number of residents in the Netherlands, we obtain the relative population density (last column in Table 4). This ratio is taken as input for the population density in the model. Data is gathered from the online topographical database Metatopos (Metatopos, 2020).

Table 4 Population density and spread per province

Province	**Population density (residents per km²)**	**Relative population density**
Groningen	252	0.5
Friesland	195	0.4
Drenthe	188	0.4
Overijssel	350	0.7
Flevoland	300	0.6
Gelderland	420	0.8
Utrecht	912	1.8
Noord-Holland	1081	2.1
Zuid-Holland	1374	2.7
Zeeland	215	0.4
Noord-Brabant	523	1
Limburg	520	1

In the table in Appendix C, we adjust the share of infections, hospitalizations, and deaths for population density. This means that we take the real number of infections, hospitalizations, and deaths for each province and adjust all relative population densities to 1. In this situation all provinces have the same population density. We observe that after adjusting for population density there remain considerable differences in the percentage of infections, hospitalizations, and deaths per province. Noord-Holland and Zuid-Holland show by far the highest share in infections before adjusting for population density. The share of hospitalizations and deaths in these two provinces are also high compared to other provinces. After adjusting, Gelderland and Noord-Brabant show the highest share in infections, hospitalizations, and deaths in the Netherlands. What we learn from this is that population density influences spread but does not fully explain differences in spread between provinces. Differences in other factors (e.g. contact density, age, and hospital capacity) are also causing differences in spread between provinces. The population density will not be included in further analysis, because this factor cannot be influenced in reality and the main focus of this study is to analyse spread of COVID-19 in the Netherlands.

Adoption of government measures

It is hard to quantify the extent by which people adopt government measures. We take the adoption of government measures into account in two ways. We use the number of people staying at home per day (staying at home behaviour) and the extent by which people travel with public transport per day (traveling behaviour). Google provides data about how often people stay at home per day compared to a baseline. We quantify traveling behaviour with data about public transport.

Staying at home behaviour

In Figure 19, data from Google about mobility in the Netherlands is visualised from 10 March (just before implementation of the first government measures) until 31 December (Google.nl, 2020). The higher the observed value, the more people stay at home compared to the baseline. As a baseline, Google uses data from January 2020, before the implementation of measures in the Netherlands. We observe seasonality within a week, with lower percentages in weekends than on other days. This is assumed to be caused by people staying at home in weekends more often than on other days. To remove the effect of weekly seasonality, we additionally show the average percentage of people staying home per week.

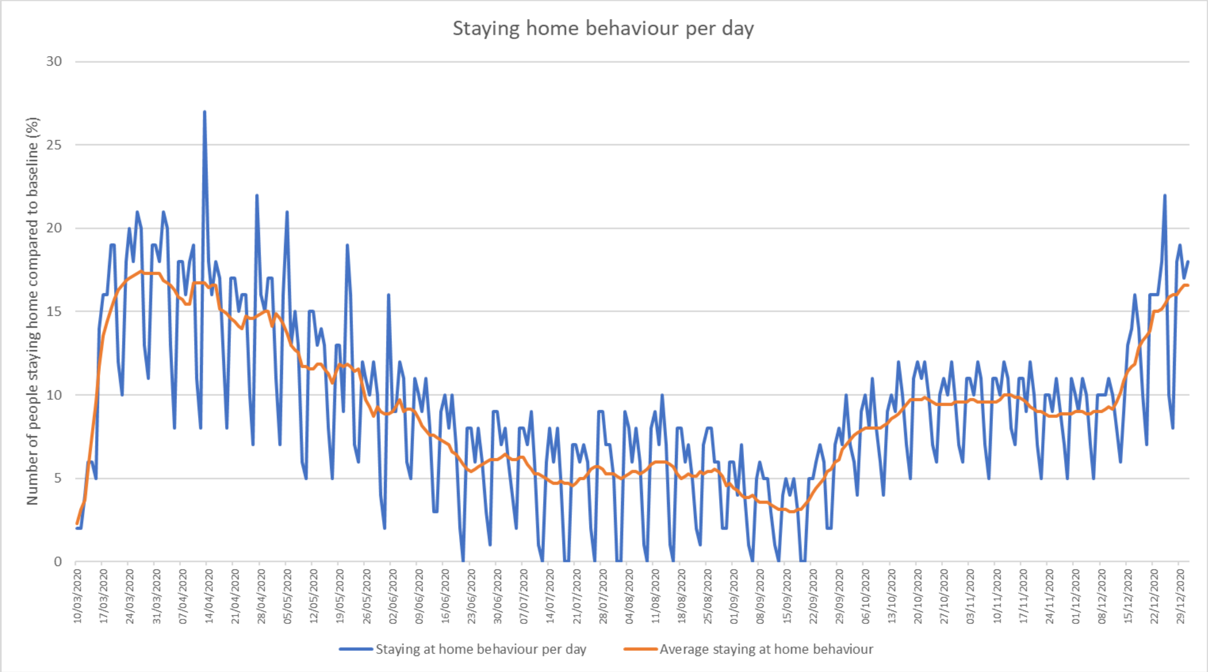


Figure 19 Staying at home behaviour in the Netherlands

After first measures were implemented on 12 March we see a fast increase in the percentage of people staying at home compared to baseline, to approximately 18%. This percentage gradually declines to approximately 4.5% until around 18 September. Around this time, the government implemented stricter measures to prevent spread of the virus. This caused the number of people staying at home to gradually increase after this date. On 15 December, lockdown measures were implemented for the second time and the number of people staying home rises quickly to a similar level as at the start of first lockdown measures. We learn from this that changes in implemented measures have significant impact on staying home behaviour. The exact effect will be identified with help of statistical analysis.

Traveling behaviour

In Figure 20, we provide data from Translink about the number of check-ins per day in public transport for 10 March until 31 December (Translink, 2020). As a baseline we use the number of check-ins on 8 March to represent the number of check-ins per day before any measures were implemented. The number of check-ins on other days is taken relatively to this baseline. We observe the same weekly seasonality as with staying home behaviour. We assume this again to be caused by the fact that, before the start of the pandemic, people use public transport less often in weekends than on other days of the week. To remove the effect of seasonality, we take the average number of check-ins per week to represent values per day.

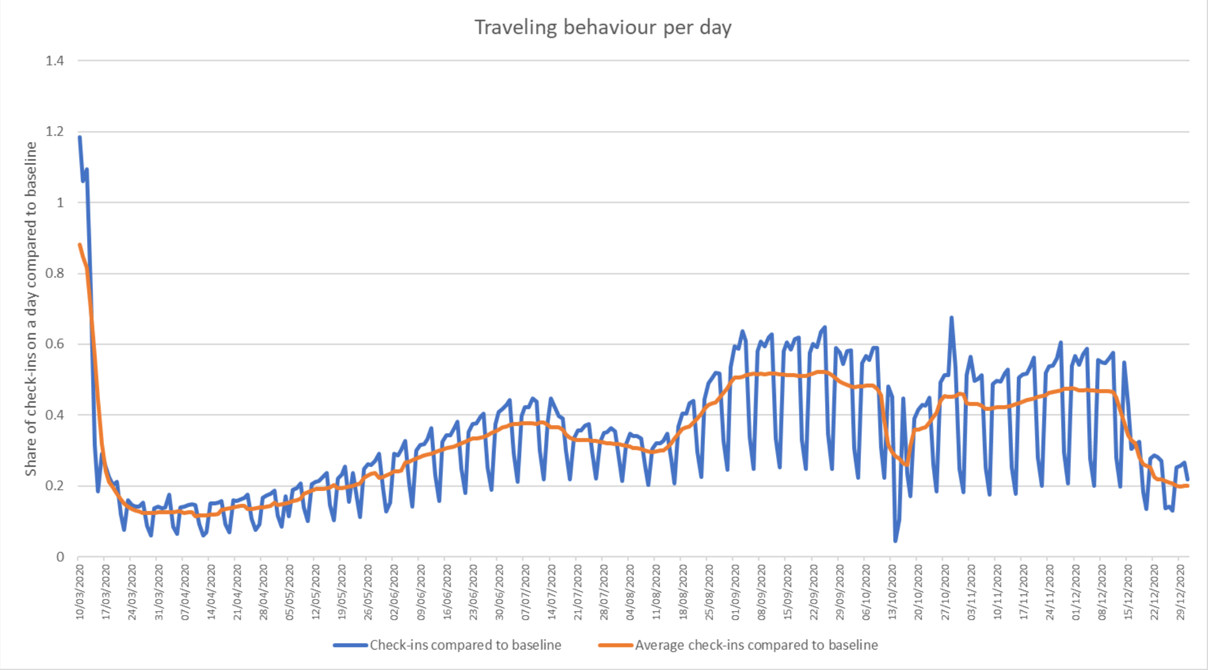


Figure 20 Traveling behaviour in the Netherlands

When first government measures were implemented on 12 March, the number of check-ins declined fast. Thereafter, the number of check-ins per day gradually increased until 8 July. On this date summer holidays started, which means people travel to school or work less often. After 8 July we see a small decline and around 17 August, when first schools opened again, the number of check-ins increased. On 1 September all schools started again and the number of check-ins per day remained rather equal until around 10 December.

On 28 September and 13 October, stricter measures were introduced. On both dates we see a decrease in traveling behaviour, of which a considerable decrease on 13 October. Around one month later, traveling behaviour had increased again to almost the level before 13 October. Traveling behaviour remained equal for quite some time thereafter. On 15 December, new lockdown measures made traveling behaviour decrease to 0.2. Based on these observations we believe that adjusting of measures affects traveling behaviour, yet effects do not always last long. What we learn from this is that there are considerable changes in traveling behaviour that can explain the adoption of government measures.

Places of infection

Another sub-factor we use to express contact rate is with places where people are exposed to infection, which we refer to as places of infection. In some places, the number of infections is observed to be much higher than in other places. Most government measures are implemented to reduce contact rate in places of infection. The intelligent lockdown implemented in the Netherlands to prevent spread during the first peak tempered contact rate a lot. Effects of this lockdown were severe, both negatively and positively. Spread of the virus was mitigated, but economy suffered (see Chapters 1 and 2). After a while social distancing measures were gradually softened and facilities (e.g. schools, restaurants) partially opened again, leading to an increasing contact rate. At the end of August the increasing contact rate lead to a new peak in number of infections (de Beus, Rooze, & van Bree, 2020). Without the implementation of a second lockdown, economic consequences were expected to remain absent (NU.nl, 2020e). Unfortunately, the Dutch government was forced to implement new lockdown measures to prevent further spread of the virus at the middle of December.

In Table 5, the percentage of infections in a certain place is provided. Data in this table is based on the source- and contact study of the GGD. Unfortunately, in approximately 60% of the cases the exact place of infection is hard to identify and thus not known (van Kempen & Soetenhorst, 2020). The numbers in Table 5 are therefore only an approximation. Data in the table shows the average percentage for the period from 31 August until 12 January (RIVM, 2020j).

Table 5 Percentage of infections per place of infection reported to the GGD

Place of infection	**Percentage of infections**
At home	52.7
Visit at home (friends/family)	22.1
Work	14.2
School	5.5
Traveling/holidays	0.3
Catering services	0.6
Party/birthday	0.4

By more than half of the infected cases, “at home” is reported as a possible place of infection. Thereafter, “work” and “visit at home” are reported as most probable place of infection with a total of 36.3% of reported places. School, catering services, traveling/holidays and party/birthday are reported by around 5% or less than 5% of the cases. The number of reported possible infections in remaining places are low and therefore not considered separately.

From all reported places of infection in Table 5, we consider “school” and “catering services” as places that can have considerable influence on spread and are thus sensitive for measures. We believe that the number of reported places of infection for catering services is low (0.6%), because catering services were closed before 1 July and after 31 August. We include the place “events”, which represents gatherings with more than 30 people (e.g. festivals). Big events were prohibited from 15 March to 31 December and thus not reported to the GGD as possible place of infection. However, due to the big impact this factor might have on spread, we include this factor as a place of infection in the model. The places at home, visit at home, work, traveling/holidays, are taken into account with the sub-factors staying at home behaviour and traveling behaviour that are discussed on the previous pages. In the remainder of this report, we refer to included places of infection as school openings, catering service openings, and event allowances. In Section 3.3, we quantify these places of infection based on implementation of measures concerning school openings, catering services openings, and events. Thereafter we can determine the effect of resulting measures on spread, which we do in Section 4.3.

Factor 8: Testing capacity

In Section 2.2, we mention the effect of a change in test availability on the development of indicators. We assume that testing capacity can affect the number of cases in two ways. The first is that testing capacity can influence the total number of infections. The second is that testing capacity can cause a discrepancy between the number of infected cases and the number of confirmed cases. Below, we outline the development of testing capacity in the Netherlands, we provide the number of tests per week, we discuss testing and reporting delay, and differences in testing capacity between provinces.

Development of the testing capacity

At the start of the pandemic, testing capacity was minimal. Only people in risk groups, patients with high risk on a severe disease and healthcare personnel, could apply for a test. Testing capacity first increased on 6 April. This expansion of test capacity was only meant for workers in healthcare and for patients with high risk on severe COVID-19 disease (RIVM, 2020k).

From 1 June, everybody with mild symptoms could apply for a test. The government assumed that in the first week of June the number of requests would be very high, causing the GGD to need more time for testing (Rijksoverheid, 2020e). From the 500,000 calls the GGD received in the first week, only a small part was performed. The RIVM calculated that 30,000 tests would have to be performed per day to be able to treat all people with complaints. But at that moment, pressure could not be handled (van de Klundert, 2020).

On 12 July the media company NU.nl wrote about an increase in the number of requests. At that moment GGD’s could still handle the pressure, but it was not clear for how long. At 28 July, test capacity was almost reached with an increase in the number of requests of approximately 1,000 per day, fortunately still sufficient. While testing capacity was sufficient, it remained important to increase testing capacity to keep up the increasing number of requests (NU.nl, 2020a).

From the beginning of September, testing capacity really became insufficient when the number of infections started to increase. It could take weeks before testing capacity would be sufficient, according to an article of NU.nl (NU.nl, 2020d). In another article published on 22 September, the GGD-president declared that from the 38,000 daily requests only 28,000 could be satisfied. In the same article, the expectation was made that only from October all daily requests could be satisfied. This means that 50,000 tests would have to be available by then (NOS, 2020). Additionally, a thread developed for a new peak that could be higher than the first peak. This thread could be even higher when testing capacity is insufficient (NU.nl, 2020c). From October, testing capacity increased and testing delay decreased. Part of this increase in capacity is caused by the rise of so-called “quick tests” (Dutch: Sneltesten), which could help tracking of infected cases. In our model, we do not take quick tests into account and we assume testing capacity to be sufficient from October.

Testing and reporting delay

Test requests are mostly met within 48 hours. The outcome of a test is clear within 24 to 48 hours after taking the test. This means that after someone thinks they show symptoms of the disease, this person knows whether this is caused by COVID-19 between 48 and 96 hours (RIVM, 2020l). Testing and reporting delay influence the time between infection and confirmation, and consequently can affect spread of the virus. With a limiting testing capacity, people have to wait longer for a test, which can limit the isolation of infectious population and increase the number of infections.

Number of tests

In Figure 21, we provide the number of tests per day. Data before 1 June was gathered from Datagraver (Datagraver, 2020). Data after 1 June was gathered from the Dutch central government (Rijksoverheid, 2020a). We observe a gradual increase in number of performed tests until the middle of October, with a temporary decrease in the month November. This decrease happened just after the second peak and just before the third peak. We believe that a reduction of infected cases per day caused this temporary decline in number of tests. The number of tests per day in December is the highest of the whole period.

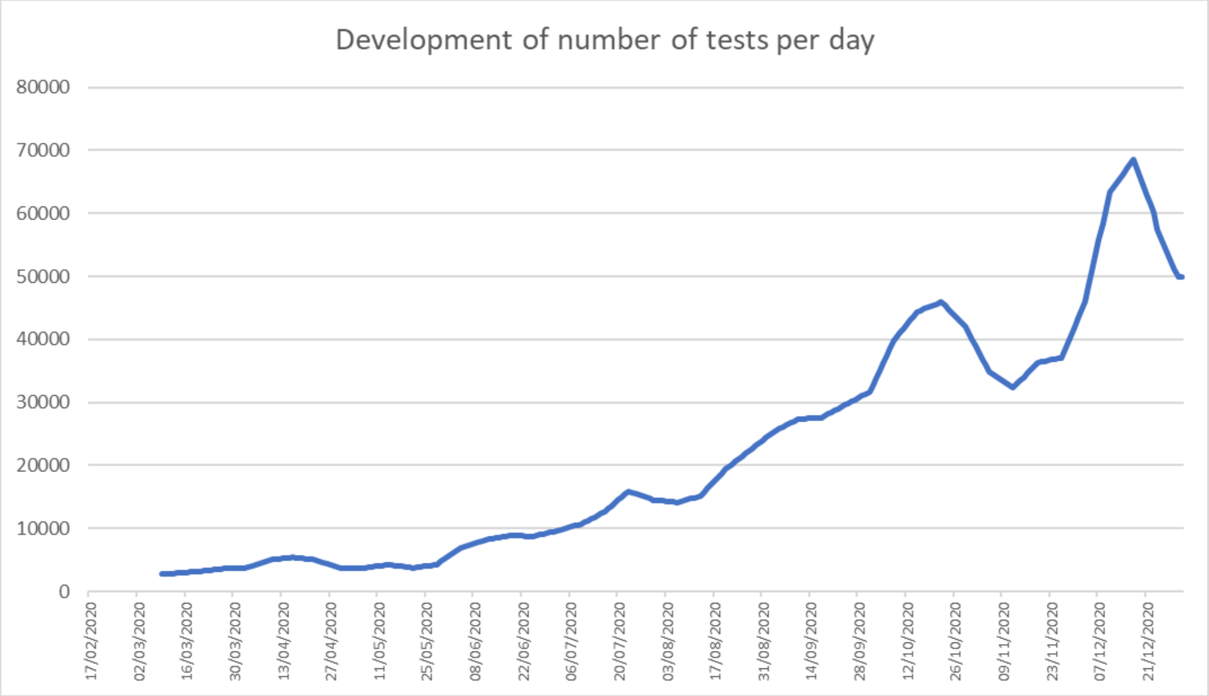


Figure 21 Number of tests per day in the Netherlands

We conclude that the number of tests increased significantly between March and December. When the number of tests in March would have been equal to the number of tests in December, the number of confirmed cases in March would have been much higher. Since we think there might exist a relation between the number of tests on a day and the reproduction rate, we study this relation in the statistical analysis. Besides, we use the number of tests to develop a regression model for the number of confirmed cases on a day in Section 4.3.4.

Differences between provinces

The place of residence determines how fast one can be tested, according to a podcast of NU.nl published on 28 July. Northern provinces like Overijssel, Friesland, and Groningen, can fulfil demand at the same day, but in a southern province (e.g. Brabant) the test has to be taken the next day mainly due to higher demand (NU.nl, 2020b). Around that time, the number of requests in the north of the Netherlands was lower than in the South. This caused testing capacity to be insufficient in nine regions. When testing capacity began to lack, it was lacking first at regions with the highest number of infections. After some time, the pressure could be partly solved when the regions with sufficient capacity replaced their capacity to regions in need of more capacity (Feenstra & Effting, 2020). Testing capacity within the whole country began to become insufficient from the first week of September (Dupuy, 2020). On 10 September, only four of hundred GGD locations had enough testing capacity, all other locations experienced pressure to deal with demand. Differences in testing capacity within the Netherlands will only be included in further analysis when we consider differences between provinces.

3.3 Measures to consider in the model

The Dutch government implemented several measures to control spread of the virus. However, the effect of applied measures is not fully understood. In our study we consider the measures facemasks and self-quarantine, and the measures that express contact rate (i.e. school openings, catering service openings, and event allowance). In Section 3.3.1, we study the effectiveness facemasks, and in Section 3.3.2 we identify the effectiveness of self-quarantine. In Section 3.3.3, we quantify the measures school openings, catering services openings, event allowance, and facemasks, to be able to determine their impact on the effective reproduction rate in the next chapter. We assume the measure self-quarantine to have a constant value in our model.

3.3.1 Facemasks

In Chapter 1 we found that the virus spreads primarily via aerosols and respiratory droplets. Respiratory droplets are more commonly causing spread according to the WHO. These droplets can spread when an infected person sneezes or coughs. Facemasks do not prevent an individual from getting infected, but might prevent an infected person from spreading the virus. Wearing facemasks everywhere outside is not considered to prevent spread. Yet in public transport it might be hard to keep 1.5 meters distance. Therefore, the use of non-medical facemasks became mandatory in public transport on 1 June. Non-medical facemasks do not satisfy the guidelines for use in healthcare, and can therefore not be used there. Medical facemasks are mandatory in healthcare. We do not take these into account in the model. Public transport was the only setting where non-medical facemasks were introduced until 20 September. After 1 October, the government strongly urged people to wear a facemask in public places. The advice to wear a facemask was already given in a press conference in the middle of September in a few regions in bigger cities in the Netherlands. Two weeks thereafter, the prime minister changed that advice to a strong urge of wearing facemasks in all public places. On 1 December, facemasks became mandatory in all public spaces.

Since there is no strong evidence for the working of non-medical facemasks, we test the effect of this measure with our system dynamics model. To be able to do this, we identify the relation between wearing facemasks and the effective reproduction rate in the statistical analysis (Chapter 4).

3.3.2 Self-quarantine

When a person is infected or waits to be tested, the person has to stay at home until the disease is gone or the outcome of the test is negative. This measure is called self-quarantine. Self-quarantine reduces the probability that an infected person spreads the virus. This measure is very effective, since it can reduce contact rate significantly. The reason for this is that the number of infectious people is based on the infectious period and the number of infections on a day (RIVM, 2020b). Shortening of the (effective) infectious period can decrease the number of infectious people. In our model, we assume that confirmed infected cases can enter self-quarantine and non-confirmed cases cannot. When a confirmed case has entered self-quarantine, the case cannot infect another anymore. We indicate the fraction of confirmed cases effectively entering self-quarantine with the “quarantine fraction”. We did not find evidence for the effectiveness of self-quarantine in literature. Therefore, we initially assume a proportion of 70% of the cases (effectively) entering self-quarantine in our model. This value might have to be calibrated.

3.3.3 Quantification of measures

The degree by which schools and catering services are opened, events are allowed, and facemasks are implemented will be quantified per day, based on the definition of measures in Table 6. In Chapter 4, assigned values per day are used to identify the effect of school openings, catering service openings, event allowances, and facemasks on the effective reproduction rate. For example, there is no contact between people in schools when all schools are closed, which can affect the reproduction rate. In the system dynamics model we can change values of measures, such that we are able to identify the impact of adjustments in measures on spread. Historical values of these measures are clarified in Appendix D, with additional dates of implementation.

Table 6 Specification of values for measures

Measure	Value	Definition
Schools openings	1	Normal education. Fully physical.
	2	Mostly physical education / Most schools open.
	3	Partially physical, partially non-physical education / Approx. half of schools opened.
	4	Mostly non-physical education / most schools closed.
	5	All schools closed / only online education.
Catering service openings	1	All catering services normally opened.
	2	All catering services normally opened with distancing measures.
	3	All catering services opened, closing at 12 PM. With distancing measures
	4	All catering services opened, closing at 10 PM. With distancing measures
	5	All catering services closed.
Events allowance	1	All events allowed.
	2	Some events cancelled.
	3	Big events prohibited.
	4	Events prohibited and gatherings with maximum number of people.
	5	All events and gatherings are prohibited. No maximum
Facemasks	1	No facemasks.
	2	Facemasks in public transport only.
	3	Urge to wear facemasks.
	4	Strong urge to wear facemasks in all public spaces.
	5	Facemasks mandatory in all public spaces.

3.4 The model

Compartment models in which individuals are passing through several stages of a disease are a common way for epidemiology models to base outbreak predictions on, for example the number of infections. The model of Kermack and McKendrick, the so-called SIR-model, is a well-known epidemiological model that can be used to determine the number of infections. Extensions of this model are the SEIR and the SEIIR model. The SEIR includes a distinction between infectious and non-infectious incubation period, and the SEIIR includes a distinction between infectious and non-infectious incubation period and between asymptomatic and symptomatic cases.

3.4.1 The SIR model

This model distinguishes susceptible, infected, and recovered people in the population. The differential equations for the susceptible, infected, and recovered population are expressed with $\dot{S}$, $\dot{I}$ and $\dot{R}$ respectively.

$$\begin{aligned} \dot{S} &= -\beta SI \\ \dot{I} &= \beta SI - \gamma I \\ \dot{R} &= \gamma I \end{aligned}$$

Transition rates between different states are inverses of the contact period B = 1/β and the infectious period C = 1/γ. Fractions S, I, and R represent the fraction of susceptible, infected, and recovered individuals respectively (Peirlinck et al., 2020). We assume the infectious period to be 9 days (see Section 3.2) and thus obtain a transition rate $\gamma = \frac{1}{9}$. To obtain the transition rate β on day t, and consequently the number of infections on a day, we divide the effective reproduction rate $R_e(t)$ with the infectious period. A regression model for $R_e(t)$ will be developed with help of a multiple regression analysis in Section 4.3 to be able to express $\beta(t)$.

$$\beta(t) = \frac{R_e(t)}{\gamma}$$

However, in our model we want to include the fact that an individual is infected but has not yet entered the infectious period. This way we can obtain a more accurate prediction of the number of infections in the case of COVID-19. The SEIR model includes this distinction.

3.4.2 The SEIR model

A popular model to simulate the spread of COVID-19 is the SEIR model. This model adds an exposed part of the population to the SIR model, which is useful due to the relatively long latency phase of the disease. With this model we can include the fact that infected cases are not yet in their infectious period. The differential equation for each compartment is:

$$\begin{aligned} \dot{S} &= -\beta SI \\ \dot{E} &= \beta SI - \propto E \\ \dot{I} &= \propto E - \gamma I \\ \dot{R} &= \gamma I \end{aligned}$$

Here the transition rate from exposed to infectious, expressed with the inverse of the latent (pre-symptomatic) period A = 1/α, and the fraction of exposed individuals E are added (Peirlinck et al., 2020). Yet an additional extension is required when we want to be able to include differences between asymptomatic and symptomatic cases, which is included in the SEIIR model.

3.4.3 The SEIIR model

As a fifth compartment, an asymptomatic population can be included to make a clear distinction between symptomatic and asymptomatic cases. This was done in the study of Peirlinck et al. (2020). The differential equation for each compartment is:

$$\begin{gathered}\dot{S} = -S[\beta_s I_s + \beta_a I_a] \\ \dot{E} = -S[\beta_s I_s + \beta_a I_a] - \propto E \\ \dot{I_s} = v_s \propto E - \gamma_s I_s \\ \dot{I_a} = v_a \propto E - \gamma_a I_a \\ \dot{R} = \gamma_s I_s + \gamma_a I_a\end{gathered}$$

In this model, the fraction of susceptible cases is divided in symptomatic (v_s) and asymptomatic (v_a), a distinction is made between contact rate of symptomatic cases (β_s) and contact rate of asymptomatic cases (β_a), and a distinction is made between the infectious period of symptomatic cases (γ_s) and asymptomatic cases (γ_a). This leads to a total contact rate β and infectious rate γ:

$$\begin{gathered}\beta = v_s\beta_s + v_a\beta_a \\ \gamma = v_s\gamma_s + v_a\gamma_a\end{gathered}$$

3.4.4 The system dynamics model

We will model the spread of COVID-19 based on the theories of the SEIR and the SEIIR model. In first instance, we determine the spread considering the same contact rate for symptomatic and asymptomatic cases and thus model spread with the SEIR model. In the sensitivity analysis we consider differences between symptomatic and asymptomatic cases and model according to the SEIIR model.

3.5 Conclusion

Now we answer research question 2 “What factors and measures should be considered in the model?”:

3.5.1 Factors

Factors that we study in the literature review are: Incubation period, infectious period, reproduction rate, fatality ratio, age, weather, contact rate, and testing. The disease factors (incubation period, infectious period, reproduction rate, case fatality ratio) cannot be changed or conform in association with other factors. The incubation period is estimated to be between 2 and 14 days with an average of 5 to 6 days. There is not enough information about the infectious period yet. In general it is assumed that an infected person can infect other people when the infected person is showing symptoms. However, there is evidence for pre-symptomatic transmissions, in one study even estimated to be a proportion of 44% of total infections. We assume the duration of the infectious period depends on the severity of the disease. The infectious period of asymptomatic cases is assumed to last on average 6 days and the infectious period of symptomatic cases 11 days. When assuming an asymptomatic fraction of 40%, this leads to an average infectious period of approximately 9 days. The reproduction rate is the rate by which the virus spread. It indicates how many people are infected by one infected person on average. The basic reproduction rate is assumed to be 2.5. The basic reproduction rate is influenced by external factors like government measures, leading to an effective reproduction rate. The effective reproduction rate will be determined in Chapter 4. The last disease factor we study is the fatality ratio. The CFR is the ratio of confirmed cases to deaths and is estimated to be 0.9%. We use the IFR to determine the number of deaths in the system dynamics model, which is assumed to be between 0.3% and 1.0% and initially included as 0.9% in the model.

The age of an individual influences the probability of infection, hospitalization and death. We believe provinces experienced differences in spread which are caused by differences in age-distribution. In the first peak of infections, the number of hospitalizations and deaths was higher than in the second peak. At the same time, average age of confirmed cases was considerably lower in the second peak. Similar to the disease factors, age of individuals cannot be influenced with government measures. The impact of age will thus not be included in our model.

The next factor considered to influence spread of COVID-19 is weather. Values of this factor cannot be influenced with government measures. To express weather, we study the sub-factors temperature, humidity and wind speed. Higher temperature shows to have a reducing effect on spread according to the literature review. The relation to spread of other factors is not clear from the literature study, we determine the relation between spread and the sub-factors of weather with help of statistical analysis.

Next we study the effect of contact rate on spread. The goal of most government measures is to reduce contact rate, which makes it important to study contact rate when studying the effect of measures. We use sub-factors population density, adoption of government measures, and places of infection to express contact rate. Higher contact rates are observed in provinces with higher population density. Besides, contact rate increases when people take government measures less serious. We express the extent by which people adopt government measures with staying home behaviour and traveling behaviour. Places of infection we identify to be important for spread are schools, catering services, and events. Effects of these places of infection are taken into account with the measures school openings, catering services openings, and event allowance.

Next to contact rate, effects of testing are large. In periods with low testing capacity (e.g. from February to June), the number of confirmed cases was far less than the number of infected cases. When people have to wait for a test in periods with an insufficient testing capacity, it is more likely that isolation of infectious cases is limited. This can increase the number of infections even further.

3.5.2 Measures

In this study, we consider the measures school openings, catering services openings, event allowance, facemasks, and self-quarantine. We quantify the measures school openings, catering services openings, event allowance, and facemasks per day with a value 1, 2, 3, 4 and 5, indicating strictness of the measure (see Table 6). When adjusting measures later in this study, the values of these sub-factors can be tuned according to the specification of values in Table 6. Effectiveness of the measure self-quarantine on spread is not clear (yet). We assume the quarantine fraction, the fraction of confirmed cases effectively entering self-quarantine, to be 70%. This value might have to be calibrated to obtain actual spread, which we do in the next chapter.

3.5.3 The system dynamics model

In first instance we built a SEIR model, meaning that we do not differentiate between symptomatic and asymptomatic cases. The SEIIR model will be applied in the sensitivity analysis when differences between asymptomatic and symptomatic infections are considered.

4. Method

This chapter is especially important for the development of a reliable model. First the study design will be outlined in Section 4.1. In Section 4.2, we estimate values of factors when necessary, and in Section 4.3, we identify relations between parameters with help of statistical analysis. In Section 4.4, parameter values are eventually calibrated to mimic actual spread.

4.1 Study design

We develop two regression models that we use in our system dynamics model, to obtain a model that can predict spread of COVID-19. One regression model expresses the effective reproduction rate and the other expresses the number of confirmed cases on a day. The approach we use to develop the regression models and the system dynamics model is provided in Sections 4.1.1 and 4.1.2 respectively.

4.1.1 Developing the regression model

The aim of this study is to predict the effect of factors and measures on spread of COVID-19 in the Netherlands. Factors that are tested for their relevance are "other factors" (temperature, humidity, wind speed, staying home behaviour, traveling behaviour, and testing capacity) from Section 3.1. Measures that are tested for their relevance are school openings, catering service openings, event allowance, and facemasks. In Figure 22, a flow chart visualizes how we obtain relevant input parameters, which we do in Section 4.3.

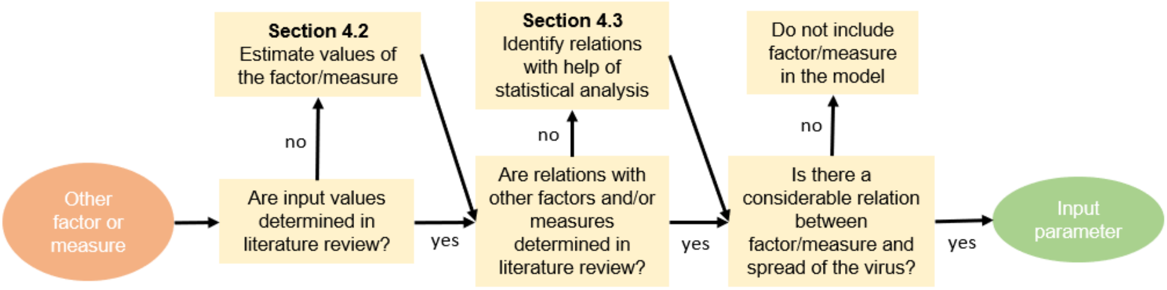


Figure 22 Approach to identify relevant input parameters for the regression model

Factors and measures are considered to be relevant parameters when they show a statistically significant relation to the effective reproduction rate. Together these relevant parameters are used to predict the effective reproduction rate with a regression model. Next to the reproduction rate, we develop a regression model for the number of confirmed cases per day. This will be done with the number of tests and the number of infected cases on a day.

4.1.2 Developing and calibrating the system dynamics model

With the regression models for the effective reproduction rate and the number of confirmed cases on a day, we are able to develop the system dynamics model in Section 4.4. Next to the relevant factors and measures in these regression models, disease factors have to be implemented in the system dynamics model. We use fixed input values for disease factors (i.e. incubation period, infectious period, basic reproduction rate, and fatality ratio), based on estimates from the literature review. These values will eventually be calibrated to mimic spread. Besides the disease factors, the initial number of infected cases and quarantine fraction might have to be calibrated to obtain actual spread. In Figure 23, a flow chart visualizes the approach we apply to calibrate parameters.

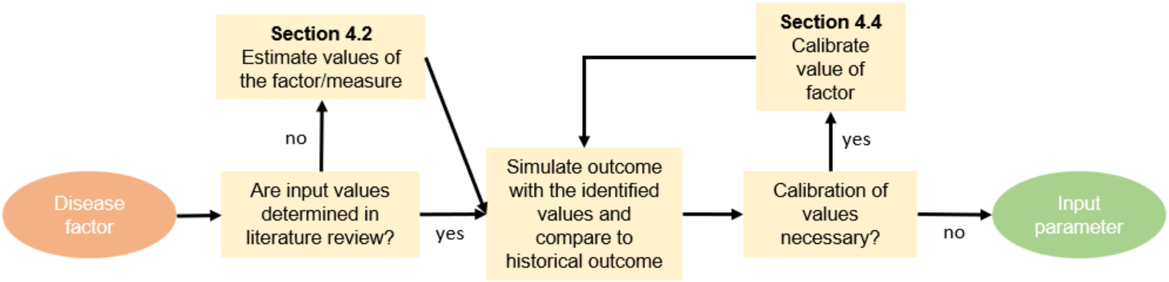


Figure 23 Approach to calibrate values of input parameters in the system dynamics model

Before we can calibrate the model, we have to know the values of indicators for the simulated period. Indicators serve as output of the model, they indicate severity of the pandemic. In Section 4.4, we provide actual values of indicators on 31 November, since values until 31 November are used to train the model. After calibration, we still have to validate the model before we can use it. This will be done in Chapter 5 with help of data from 1 December until 31 December.

4.2 Estimating values of factors

Values of the factors incubation period, infectious period, basic reproduction rate, fatality ratio, temperature, humidity, wind speed, staying home behaviour, traveling behaviour, and number of tests, and of the measures school openings, catering service openings, event allowance, and facemasks have been determined in the literature review. Input values of these parameters are provided in Table 14. Only the effective reproduction rate has to be estimated.

4.2.1 Estimating the effective reproduction rate

The relation between the effective reproduction rate $R_e(t)$ and all relevant parameters will be determined in Section 4.3 with help of multiple linear regression. Yet before we can identify these relations, we have to calculate the actual $R_e(t)$, to be able to compare the outcome of our model to the actual reproduction rate. We refer to the calculated effective reproduction rate as $R_e{}^{calculated}(t)$.

Historical values of the effective reproduction rate will be calculated with the number of hospitalizations or confirmed cases per day from 27 February until 31 December. $R_e{}^{calculated}(t)$ is calculated in R with the function *est.R0.TD*. This is an implementation of the method provided by Wallinga & Teunis (Wallinga & Teunis, 2004). The method estimates the time dependent reproduction number, together with a confidence interval. Before we can use this function, we have to make assumptions about the mean generation interval for COVID-19. The mean generation interval represents the average time between the infected person's first day of illness and the first day of illness for the person who infected them. The mean of this interval is approximately 4 days according to the RIVM (RIVM, 2020e). Since the RIVM does not provide an estimate for the standard deviation, we assume the standard deviation of the generation interval to be 3 days (Knight & Mishra, 2020). This means that we assume a gamma-distribution for the generation interval with a mean of 4 days and a standard deviation of 3 days in this function. A visualization of $R_e{}^{calculated}(t)$, from 20 February until 31 November, is provided in Appendix E, based on the number of infections and the number of hospitalizations respectively. We observe that the confidence interval of $R_e{}^{calculated}(t)$ based on the number of hospitalizations is larger than the confidence interval based on the number of confirmed cases. We expect this observation, since incidence of hospitalizations approaches zero around June and is much lower than the incidence of confirmed cases. When developing the regression model of $R_e(t)$ in Section 4.3, we use $R_e{}^{calculated}(t)$ based on the number of hospitalizations before 1 June and $R_e^{calculated}(t)$ based on the number of confirmed cases after 1 June as actual $R_e(t)$. When calibrating the system dynamics model, we test whether this choice might have to be reconsidered [1].

We calculate $R_e^{calculated}(t)$ for each day until 1 June based on the number of hospitalizations per day and $R_e^{calculated}(t)$ for each day thereafter based on the number of confirmed cases. We do this to increase reliability of $R_e(t)$, because the number of hospitalizations was very low after the first peak of infections and before 1 June not every infected person was allowed to be tested. This can cause a distorted approximation of the actual $R_e(t)$. To be able to determine $R_e^{calculated}(t)$ based on the number of hospitalizations, we have to know the day of infection of hospitalized cases. We assume this to be 10 days before hospitalization, including an incubation period of 5 days, testing and reporting delay of 2 days, and a time between symptom onset and hospitalization of 3 days. For the confirmed cases we estimate the day of infection to be 7 days before confirmation, including an incubation period of 5 days and a testing delay plus reporting delay of 2 days. We observe in Figure 24 that these delays provide a good fit when comparing $R_e^{calculated}(t)$ to $R_e(t)$ of the RIVM. Observed differences in fluctuations are probably caused by different input settings for the function *est.R0.TD*. The RIVM did not publish the exact assumptions they make when calculating $R_e(t)$ with this function (yet). Since we do not have sufficient data to calculate $R_{e\ calculated}(t)$ after 31 November, our calculation stops there.

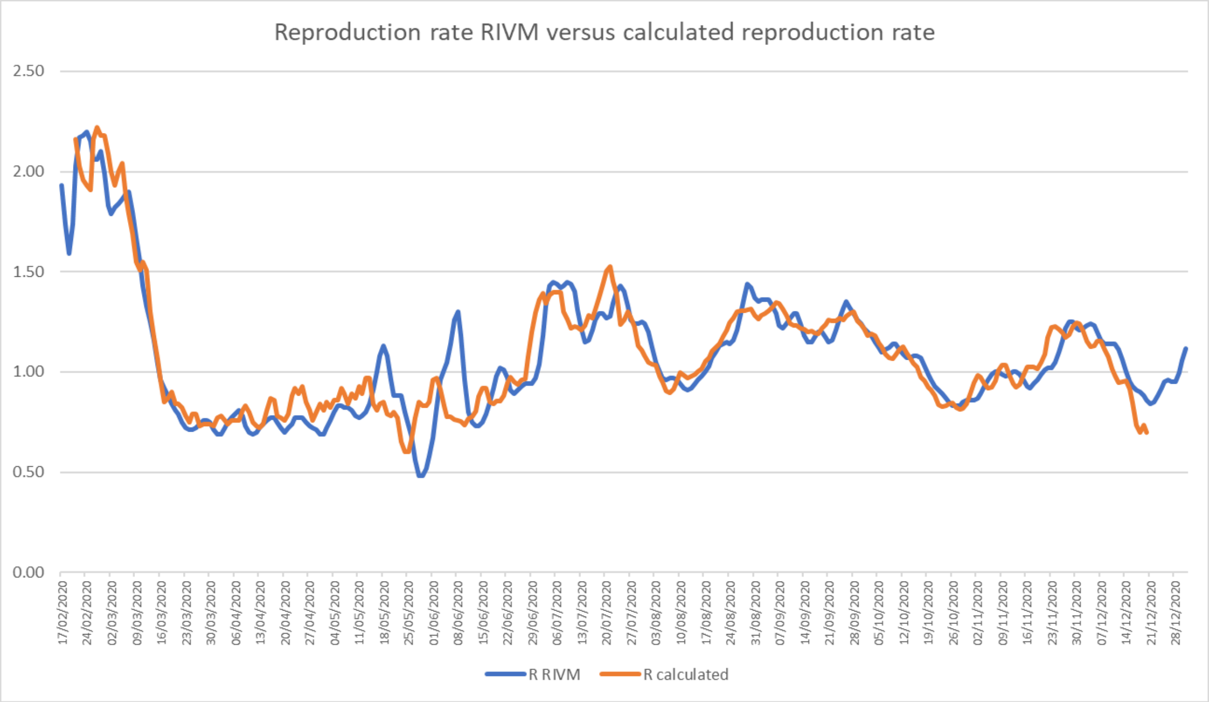


Figure 24 Comparing the calculated reproduction rate to the reproduction rate of the RIVM

---

[1] During calibration we discovered that $R_e^{calculated}(t)$ was too high after the first peak. Consequently, we were not able to obtain actual spread and had to update $R_e^{calculated}(t)$. Instead of using the average value of $R_e^{calculated}(t)$, we now use the lower bound of $R_e^{calculated}(t)$ until 15 May. From 15 May until 1 June we use the mean of the lower bound of $R_e^{calculated}(t)$ and the average of $R_e^{calculated}(t)$. From 1 June, we use the mean value of $R_e^{calculated}(t)$ based on the number of hospitalizations. This way we get a steeper decline of the infectious population after the first peak and spread in our model comes closer to actual spread.

4.3 Statistical analysis

It is possible that not all factors we study in the literature review should be included in the model. Some may not considerably influence spread and can be left out. Of key factors and measures that are influencing spread of the virus it should be clear how to include these in the regression model. Whether these parameters influence spread and to what extent will be investigated in this section. To do this, we express their effect on the effective reproduction rate ($R_e(t)$) with help of a regression model in Section 4.3.3. Next to $R_e(t)$, we develop a regression model for the number of confirmed cases per day in Section 4.3.4, to be able to estimate the number of confirmed cases per day in our system dynamics model. We provide the dataset we use and we explain the approach we apply to develop the regression models in Section 4.3.1 and Section 4.3.2. All coding in R is provide in Appendix Q.

4.3.1 Dataset for multiple linear regression

The dataset we use to estimate effects of factors and measures on $R_e(t)$ consists of 313 days of data, starting on 14 February and ending on 31 November. In Appendix F, a summary of the characteristics of all factors and measures can be found. Factors and measures that are tested on their relation to $R_e(t)$ are relative humidity, temperature, wind speed, staying home behaviour, traveling behaviour, school openings, catering service openings, event allowance, facemasks and number of tests [2]. Besides these parameters, we test whether adding the average $R_e(t)$ of two and/or one week before improves prediction accuracy. There remain 265 days of data after removing rows with missing data. Rows with missing data are mainly due to unavailable data about number of tests per day before 9 March.

4.3.2 Developing a regression model

With multiple linear regression we can identify the effect of each predictor on a response, in our case on $R_e(t)$ or on the number of confirmed cases. We can also check whether relations between predictors, factors and/or measures, exist. To identify relations between predictors, we insert interaction terms in the regression model and compare this model to the model without interaction terms. This way we can determine the added value of including interaction terms for prediction of spread. We also check whether adding non-linear variables might improve prediction of the spread.

Obtaining a multiple (non-)linear regression model

In the multiple linear regression, we assume the following formula where Y is the response and X_j are the predictors:

$$Y \approx \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \in$$

Quantifying the response with help of multiple variables is known to be a regression problem. Multiple linear regression is a relatively simple yet useful method to determine the response. The formula above is a linear model that provides an approximation of reality. In this formula, all predictors X_j are considered to have a linear effect on Y. Due to the rough estimations of some parameters (e.g. school openings), we have to accept the fact that it is not possible to reproduce the real $R_e(t)$ and number of confirmed cases using linear regression. The term $\in$ refers to a random error of Y, which is independent of the rest of the formula and has a mean of zero. Coefficient β_j is the average effect on Y when X_j increases with one unit. Non-linear variables are added to the regression formula by changing X_j to ${X_j}^i$, where i indicates the power. For example, the square root of a variable makes X_j turn into ${X_j}^2$. Interaction terms are added to the regression formula by multiplying one or more X_j's.

---

[2] Note: After some modelling we excluded the number of tests as predictor for the effective reproduction rate, because we are not able to explain the large positive effect this predictor shows. This means that we only use number of tests as predictor in the regression model of the number of confirmed cases.

Multiple linear regression can be used to determine the effect of all predictors in the model. Yet, we first need to identify the parameters that add predictive value. There exist various methods to select relevant parameters. To obtain relevant parameters for $R_e(t)$, we apply the backward regression method. The backward regression method automatically selects predictors that are considered to be important for the model. The method starts with a model that includes all predictors and one-by-one removes predictors from the model. The method stops when some stopping criteria is reached. We do this in R with the function *stepAIC*. This method keeps minimizing the per step AIC-value to develop a final set of predictors for the desired response. The AIC (Akaike Information Criteria) quantifies the amount of information that is removed in each step, as it is an estimator of out-of-sample prediction error. For the regression model of the number of confirmed cases, we do not apply backward regression. This is because we only have two predictors, the number of infections and the number of tests on a day, to develop the regression model.

Assessing the quality of the multiple linear regression outcomes

We express the statistical significance of the relation between a parameter (e.g. school openings) and a response ($R_e(t)$ or number of confirmed cases) with help of the *p*-value. We consider parameters to be statistically significant related to the response when the *p*-value is below 0.01. The * in the outcome of a model indicates the statistical significance of the *p*-value. Here "***" stands for a *p*-value <0.001, "**" for a *p*-value <0.01, "*" for a *p*-value <0.05 and "." for a p-value <0.1.

We express the quality of a regression model with the adjusted R^2. The adjusted R^2 indicates the quality of the model with a value between 0 and 1. The closer the value is to 1, the better the model. This approach other than the regular R^2 adjusts for the number of variables in the model, because the regular R^2 increases when more variables are added (James, Witten, Hastie, & Tibshirani, 2013). We use the regular R^2 to determine R^2 corrected for optimism. In the next paragraph we will outline how we correct for optimism.

Besides expressing R^2, we want the test error of our model to be as low as possible. This is important to determine the effect of using our formula on unseen data. We express the test error with the mean squared error (MSE). Knowing the test error is especially important when the ratio of parameters *p* to number of samples *n* increases, which is the case for models that include non-linear variables and/or interaction terms. In these models we expect that overfitting might play a role (James et al., 2013).

Validation of the multiple linear regression outcomes

We are aware that overfitting is a major problem in regression modelling (Steyerberg, 2009). To justify the quality and outcomes of our regression models, we validate the obtained models in two ways. First the quality of the regression models will be corrected for optimism by performing bootstrap. The bootstrap method estimates the accuracy of the method by running a certain method (in this case backward linear regression) multiple times with a different sample set (James et al., 2013). The regular R^2 and MSE after correcting for optimism will be called "corrected R^2" and "corrected MSE". We run the backward selection method 500 times to validate the quality of our model. The corrected R^2 and corrected MSE are provided in 4.3.3 and are used to compare quality of the models.

Second we test the effect of using a different response on the final regression model. We do this only for the regression model of $R_e(t)$, because in this model we do not know which predictors add value for prediction of spread. To validate the regression model of $R_e(t)$, we test the quality of the model with the predictors included in $R_e(t)$ and the number of hospitalizations or number of deaths as response. Validation is performed with these two responses, because the effective reproduction rate determines the number of infections and consequently the number of hospitalizations and deaths in our model. This validation will be done in Section 4.3.5.

4.3.3 Regression model for the effective reproduction rate

In this section, we first discuss the effect of each predictor in Appendix F on $R_e(t)$ individually. Thereafter we obtain the regression model with multiple linear regression.

Effect of individual predictors on the effective reproduction rate

In Table 7 and 8, the normalized effects of all predictors can be found in linear form. We provide normalized effects to be able to compare the effect of each predictor on the effective reproduction rate more easily. The estimate in the table indicates the value of the coefficient β_j, and the standard error indicates the standard error tells us the average amount that this estimate for β_j differs from the actual value of β_j (James et al., 2013). Thus the higher the standard error, the worse the accuracy of the estimate for β_j. Below we shortly discuss the effects of all predictors separately.

Table 7 Estimated effects of full linear regression model (normalized)

Predictor	**Estimate**	**Std. Error**	**p-value**
Intercept	1.233	0.140	2.40e-16 ***
Humidity	0.062	0.061	0.311
Average temperature	-0.319	0.077	4.90e-05 ***
Wind speed	-0.003	0.047	0.944
Staying home behaviour	-0.652	0.105	2.29e-09 ***
Traveling behaviour	0.469	0.152	0.002 **
School openings	0.269	0.058	6.88e-06 ***
Catering services	0.086	0.079	0.278
Event allowance	-0.427	0.069	2.14e-09 ***
Facemasks	0.269	0.045	6.19e-09 ***

Table 8 Estimated effects of full linear regression model with average reproduction rate of 1 & 2 weeks before (normalized)

Predictor	**Estimate**	**Std. Error**	**p-value**
Intercept	0.198	0.146	0.177
R1 week	0.916	0.073	< 2e-16 ***
R2 weeks	-0.102	0.044	0.021 *
Humidity	-0.017	0.052	0.749
Average temperature	-0.078	0.066	0.239
Wind speed	-0.024	0.038	0.522
Staying home behaviour	-0.369	0.091	7.26e-05 ***
Traveling behaviour	-0.038	0.155	0.804
School openings	0.155	0.049	0.002 **
Catering services	0.202	0.064	0.002 **
Event allowance	-0.112	0.082	0.170
Facemasks	0.133	0.038	0.001 ***

In Table 8, we show the effect of including the average effective reproduction rate of both 1 and 2 weeks before as an additional predictor. Especially the effective reproduction rate of 1 week before captures a lot of information for the regression model, because the estimate for β_j is relatively high. We call this model the full linear model of ${R_e}^{average}(t)$. The linear model without these predictors included (Table 7) is called the full linear model of ${R_e}^{linear}(t)$. In these two models, we question the estimate of β_j of predictors when the *p*-value is not statistically significant (> 0.01). As we can see, in ${R_e}^{linear}(t)$ this is the case for humidity, wind speed, and catering service openings. The estimates of β_j for these predictors is relatively small. In ${R_e}^{average}(t)$, the *p*-value is higher than 0.01 for the predictors humidity, average temperature, wind speed, traveling behaviour, and catering service openings. Since the effect of many predictors in the full linear model of ${R_e}^{average}(t)$ is not statistically significant, we only discuss effects of the predictors in ${R_e}^{linear}(t)$.

After the intercept, staying home behaviour has most effect on the effective reproduction rate in ${R_e}^{linear}(t)$. When people stay home more often, $R_e(t)$ decreases. Event allowance and traveling behaviour show most effect on the reproduction rate thereafter. Stricter measures for the allowance of events and less traveling both reduce $R_e(t)$. Also average temperature has a relatively large effect on spread, where higher temperature leads to a lower $R_e(t)$. Stricter measures for school openings, catering service openings and facemasks increase $R_e(t)$ according to the regression model. The effect of catering service openings is small and not statistically significant. We think the cause for is that catering services were closed for a long time, which might cause the model to have a hard time to determine the effect of catering service openings. This could also be the case for the measure facemasks. Facemasks were only implemented as a measure at the end of the period used to develop the regression model. We believe that both the effect of facemasks and school openings in the regression model are quite large. Humidity and wind speed have the lowest effect on $R_e(t)$. Decreasing humidity and increasing wind speed leads to a decrease of $R_e(t)$.

Regression models for the effective reproduction rate

Above we identify the individual effect of predictors in the multiple linear regression. Here we apply backward regression to remove predictors that add no or little predictive value from the full regression model. First, we perform backward regression on the model with all variables in linear form. Hereby we study the linear relation between $R_e(t)$ and the parameters in Appendix F. The resulting model is called the “linear model”. Second we perform multiple linear regression with variables in linear and non-linear form (e.g. $school\ openings^2$). We call this model the “polynomial model”. Lastly we perform multiple linear regression with interaction terms between variables. Interaction terms were included, because we think that some parameters are sensitive for interaction with other parameters. This model is called the “interaction model”. In this paragraph we only apply backward regression on ${R_e}^{linear}(t)$. We identify whether adding the average $R_e(t)$ of two and/or one week before improves prediction accuracy in the next paragraph.

Interaction terms we consider in the backward regression are:

- Humidity * Average temperature
- Staying at home behaviour * Traveling behaviour * School openings
- Staying at home behaviour * Traveling behaviour * Catering openings
- Staying at home behaviour * Traveling behaviour * Event allowance
- Staying at home behaviour * Traveling behaviour * Facemasks

Comparing the regression models

The outcomes of the backward regression for the linear model, interaction model, and polynomial model can be found in Appendix G. In the linear model, the backward regression method excludes the predictor catering services openings. But since we want to be able to express the effect of catering services openings, we add this predictor to the outcome of the backward regression model. From the interaction model, we exclude predictors that are not statistically significant relation (*p*-value > 0.01).

In Appendix H, the histogram of residuals of the linear model and the interaction model are provided. We observe the residuals to be approximately normal distributed around zero, meaning both models indicate no non-linearity.

The outcomes for the quality of all three models is provided in table 9. We observe that the polynomial model shows the highest MSE. We think this is because the polynomial model is overfitted. Meaning the model can predict $R_e(t)$ very well with current data, but not with unseen data. This would also explain the high values for R^2. Since we find it important to have a low test error and the residual plots indicate no non-linearity, we exclude the polynomial model from further analysis. The MSE of the interaction model is a little lower than the MSE of the linear model. The corrected R^2 of the interaction model is higher than the corrected R^2 of the linear model. This indicates that the interaction model provides a more accurate prediction of $R_e(t)$ than the linear model.

Table 9 Quality of regression models for the effective reproduction rate

Model	Adjusted R^2	Corrected R^2	Corrected MSE
Linear	0.791	0.783	0.0197
Polynomial	0.918	0.858	0.0319
Interaction	0.882	0.849	0.0136

The final regression model for the effective reproduction rate

When we test the interaction model with unseen data, we observe some strange effects. For example, when setting traveling behaviour equal to traveling behaviour before the start of the pandemic, the effective reproduction rate fluctuates from around 2.5 to around 0.5 and back in a few days. The same is observed when changing input values of other predictors in this model. Because of these effects, we think the interaction model is overfitted. The interaction model can provide an accurate prediction with the dataset we use to build the model, but not for unseen data. The linear model does not show these effects. For this reason, we use the linear regression model to predict spread in the system dynamics model. We refer to this model as ${R_e}^{linear}(t)$.

For the linear model, we test whether adding the average $R_e(t)$ of 1 and/or 2 weeks before to the backward regression model improves prediction accuracy. We refer to this model as ${R_e}^{average}(t)$. In Table 10 we observe that the quality of the linear model improves with the average $R_e(t)$ of 1 and/or 2 weeks before included. The adjusted R^2, corrected R^2, and the MSE of linear ${R_e}^{average}(t)$ all score better after adding these predictors.

Table 10 Quality of the linear and averaged linear regression model for the effective reproduction rate

Model	Adjusted R^2	Corrected R^2	Corrected MSE
${R_e}^{linear}(t)$	0.791	0.783	0.0197
${R_e}^{average}(t)$	0.869	0.864	0.0099

We believe the reason for the quality of linear ${R_e}^{average}(t)$ to be higher is that this regression model provides less fluctuations for the effective reproduction rate, which means we obtain a more averaged development of the effective reproduction rate. The additional variables in ${R_e}^{average}(t)$ add extra information to $R_e(t)$, since these variables can identify time-series information (e.g. a recent trend in $R_e(t)$). However, while ${R_e}^{average}(t)$ is able to include recent trends in $R_e(t)$, we observe that the average $R_e(t)$ of 1 and 2 weeks before in ${R_e}^{average}(t)$ take a lot of information that is in ${R_e}^{linear}(t)$ expressed with relevant factors or measures. Furthermore, when comparing the development of the two regression models over time, we observe that ${R_e}^{average}(t)$ needs more time to bring $R_e(t)$ below or above 1, which we believe is caused by the additional predictors. For this reason, we consider ${R_e}^{linear}(t)$ to provide a better approximation of ${R_e}^{calculated}(t)$ on unseen data, and consequently a better prediction for $R_e(t)$ in our system dynamics model than ${R_e}^{average}(t)$. For this reason, we use ${R_e}^{linear}(t)$ to express the effective reproduction rate. In Appendix G, all predictors included in this model are provided with their estimate for β_j, the standard error of the estimate for β_j, and the *p*-value. The predictor catering service openings does not show a statistically significant relation to the effective reproduction rate in this regression model. Yet we added this predictor to the linear backward regression model, because this predictor expresses the effect of a measure and we want to be able to test the effect of all measures.

The estimates for β_j in Appendix G result in the following regression model for the effective reproduction rate:

$$\begin{aligned}{R_e}^{linear}(t) = {} & 1.22 - 0.013 * \text{average temperature} - 4.179 * \text{staying home behaviour} + 0.578 * \text{traveling behaviour} \\ & + 0.066 * \text{school openings} + 0.021 * \text{catering service openings} - 0.109 * \text{event allowance} + 0.095 \\ & * \text{facemasks}\end{aligned}$$

The normalized estimate for β_j and the standard deviation of this estimate are provided in Table 11, to give insight in the relative effect each predictor on $R_e(t)$.

Table 11 Normalized effects of the final regression model for effective reproduction rate

Predictor	**Estimate**	**Std. Error**
Intercept	1.284	0.132
Average temperature	-0.339	0.075
Staying home behaviour	-0.663	0.098
Traveling behaviour	0.477	0.145
School openings	0.265	0.058
Catering services	0.085	0.079
Event allowance	-0.434	0.068
Facemasks	0.286	0.041

Of the final model, all VIF (Variance Inflation Factor) values are below 10. The VIF value can assess multi-collinearity within a regression model. In practice, there typically exists a small amount of collinearity among predictors in a regression model. A VIF value that exceeds 5 or 10 indicates an amount of collinearity that might be problematic (James et al., 2013). Only staying home behaviour, traveling behaviour and catering services openings have a VIF value above 5, yet all below 10. In Section 4.3.5, we further validate this model by assessing the quality of this regression model with the number of hospitalizations or deaths as response.

4.3.4. Regression model for the number of confirmed cases

In our system dynamics model, we have to estimate the number of confirmed cases. The number of confirmed cases highly depends on the number of infections and the number of tests per day. We perform multiple linear regression to identify the relation between these three variables. We observe a statistically significant relation between the response number of confirmed cases per day and the predictors number of tests per day and number of infections per day. For this reason, we develop a formula that predicts the number of confirmed cases on a day with help of the number of infected cases and the number of tests on a day.

The number of infected cases per day is unknown and has to be estimated before we can develop a regression model for the number of confirmed cases. We do this with data of the number of infectious cases. In Figure 25, the number of infectious cases per day is visualised. This figure originates from the weekly update of the RIVM (RIVM, 2020j). The dark purple line in the figure indicates the average infectious cases, and the light purple area indicates the lower and upper bound of the number of infectious cases. The number of infectious cases in this graph was determined with help of serological information and the number of hospital admissions. We believe the infectious population can provide an accurate indication of the number of infections when dividing the number of infectious cases on a day with the infectious period.

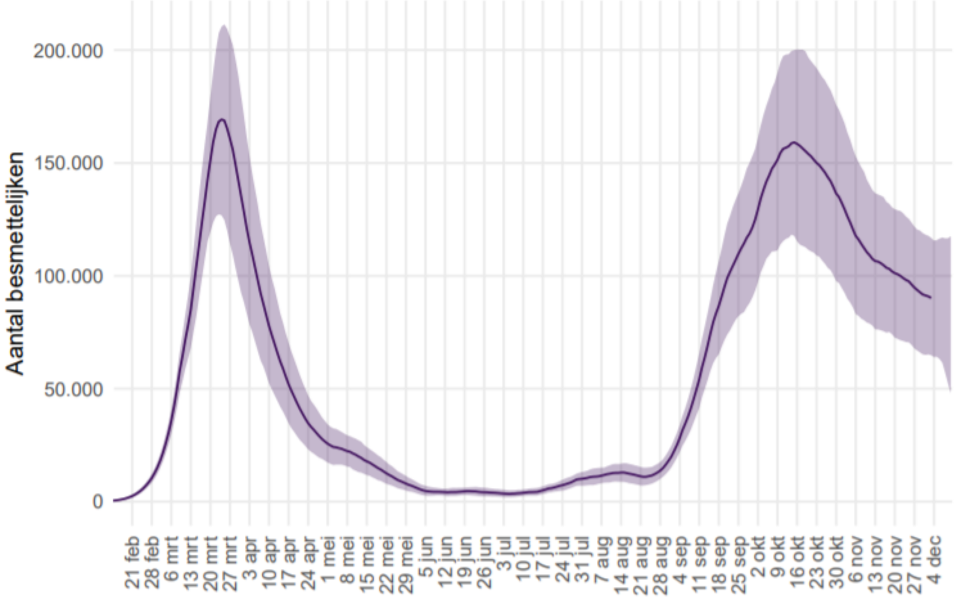


Figure 25 Development of the number of infectious cases per day according to the RIVM

The calculated number of infected cases per day according to the approach above is uncertain and can be high. This means the outcome of the linear regression model could become negative when we use different input values than the values we use to train the regression model. For this reason, we compare the linear regression model with the number of confirmed cases on a day as response and the linear model with the logarithm of the number of confirmed cases as a response. With the second model, the predicted number of confirmed cases cannot become zero. When we compare the two regression models, we observe the residuals in the residual plot to be more equally distributed in the model with a logarithmic response (see Appendix I). For this reason, we believe that the model with a logarithmic response can provide a more reliable outcome than the model without logarithmic response. This means we obtain the following formula for the number of confirmed cases on day t:

$$confirmed\ cases\ (t)\ =\ e^{4.86+4.74\mathrm{e}{-05}*\mathrm{num}ber\ of\ tests+1.329\mathrm{e}{-04}*infected\ cases}$$

In Table 12, the estimates of β_i for the predictors in this model can be found, together with their *p*-value and the standard error of the estimates for β_i.

Table 12 Estimated effects of the resulting regression model for the number of confirmed cases

Predictor	Estimate	Std. error	Pr(>|t|)
(Intercept)	4.860	0.056	< 2e-16 ***
Number of tests	4.740e-05	2.483e-06	< 2e-16 ***
Number of infected cases	1.329e-04	6.800e-06	< 2e-16 ***

4.3.5 Validation of the regression model for the effective reproduction rate

In Section 4.3.3, we apply bootstrap to validate the quality, expressed in test error and R^2, for all models. In this section, we use the number of hospitalizations and number of deaths instead of the effective reproduction rate as response, to validate $R_e{}^{linear}(t)$. We compare the corrected R^2 and corrected MSE of $R_e{}^{linear}(t)$ to the corrected R^2 and corrected MSE of the models for hospitalizations and deaths.

In Table 13, results of all three models can be found. We use ten days before hospitalization as day of infection for the number of hospitalizations and 21 days (three weeks) before death as day of infection for the number of deaths. These assumptions are based on the development of the typical progress of COVID-19 (see Figure 1).

Table 13 Quality of the validation regression models

Response	Adjusted R^2	Corrected R^2	Corrected MSE
Effective reproduction rate	0.791	0.783	0.0197
Number of hospitalizations	0.695	0.676	3139
Number of deaths	0.752	0.745	388

Based on the results in Table 13 we believe that $R_e{}^{linear}(t)$ is valid, because we observe the model for the number of hospitalized cases and the model for the number of deaths have an adjusted R^2 and a corrected R^2 close to or above 0.7. We believe these values are sufficient. The MSE of the model for the number of hospitalizations and for the model of the number of deaths are a little high. However, since we do not know the exact date of infection for hospitalizations and deaths, we consider these test errors to be acceptable.

4.4 The system dynamics model

In this section, we develop the system dynamics model we use to simulate spread of the virus. As input for the system dynamics model we use the regression models for the effective reproduction rate and the number of confirmed cases from Section 4.3, and remaining parameters in Table 14. Output of the system dynamics model is expressed with indicators of spread. As indicators we use the number of infected cases, number of confirmed cases, number of hospitalized cases, number of IC occupations, and number of deaths. Before we can use the model, we identify if and where calibration of input values is necessary in Section 4.4.1. We calibrate the model by comparing actual values to the outcome of our model.

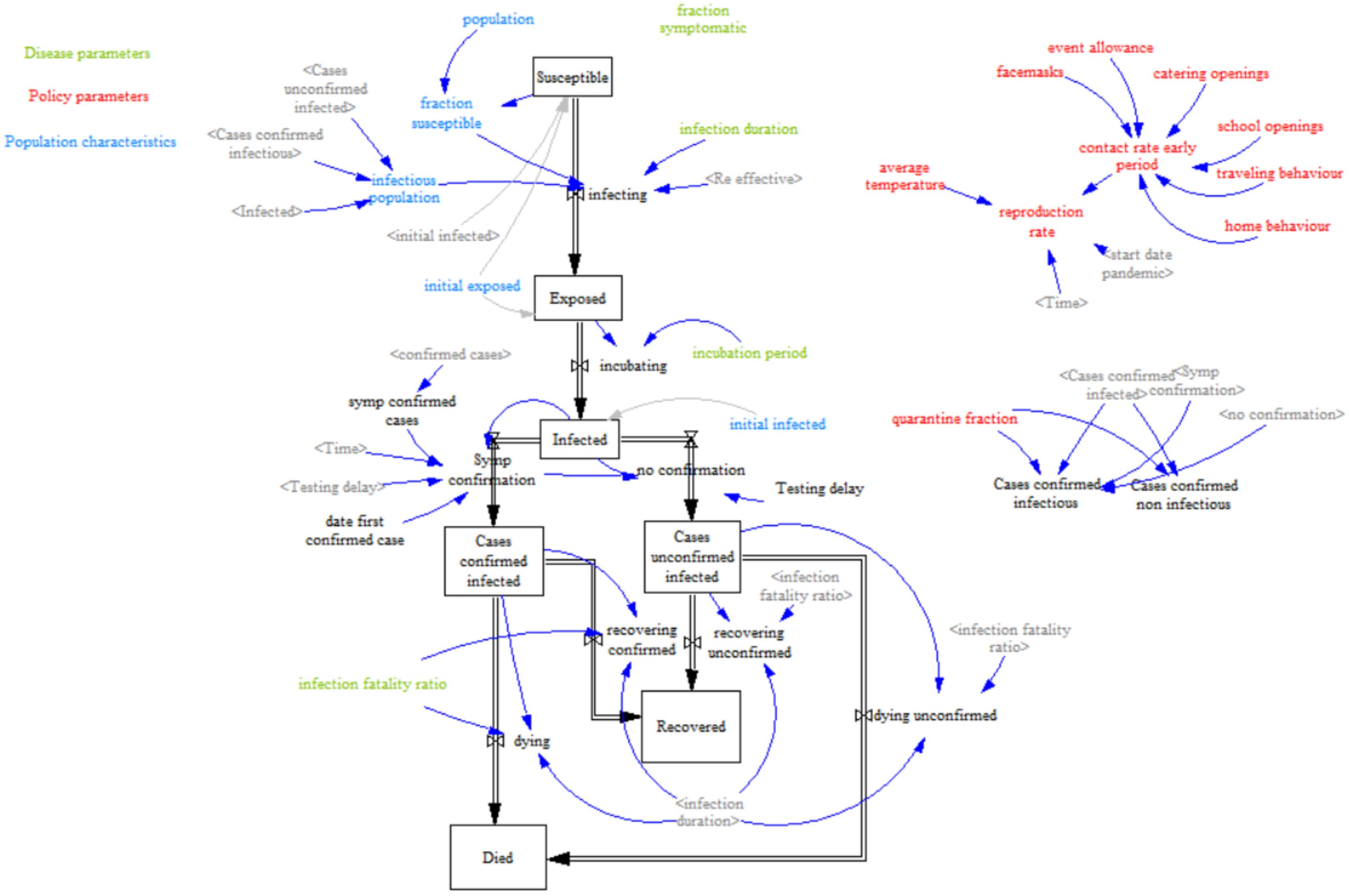


Figure 26 Visualization of system dynamics model in Vensim

In Figure 26, a visualization of our system dynamics model, developed in Vensim, can be found. This model does not differentiate between symptomatic and asymptomatic cases, which means the system dynamics model is build according to the SEIR model from Section 3.4. The squares in the model together hold all individuals in the population, each square holding a different part of the population. Individuals flow from one state to another state according to the transition rates in Section 3.4.2. Next to the susceptible population, exposed population, infected population, and recovered population, we distinguish (un-)confirmed population and died population. The (un-)confirmed population consists of infected individuals who have been tested and may consequently not be infectious anymore, when the individual entered self-quarantine. The died population consists infected cases who are not infectious anymore and did not recover but die.

4.4.1 Calibration of input parameters

We use actual values of the number of confirmed cases, number of hospitalized cases, number of IC occupations, and number of deaths on 31 November to determine whether input parameters need to be calibrated. Values are gathered from the RIVM and association NICE (National Intensive Care Evaluation), and provided in Table 15 (NICE, 2020; RIVM, 2020f). Actual values for the number of infected cases are not known and thus not provided. In Table 14, we provide calibrated values of input parameters.

Calibration values of parameters in the system dynamics model

To be able to obtain an accurate prediction of spread, we assume the first cases in the Netherlands have been infected on 2 February. We use ${R_e}^{linear}(t)$ to determine the effective reproduction rate starting from 27 February, which is the day of the first confirmed case according to the RIVM. Before this day, we assume a basic reproduction rate of 2.5 (before calibration).

Table 14 Initial and calibrated values of input parameters for the system dynamics model

Parameter	**Values**	**Value range**	**Initial input value**	**Calibrated input value**
Incubation period	*Days*	*2 – 14*	*5.5*	*5*
Infectious period	*Days*	*6 – 15*	*9*	*6*
Basic reproduction rate	*Numeric*	*0 - 5*	*2.5*	*3*
Initial infected	*Numeric*	*-*	*111*	*600*
Infection fatality ratio	*Numeric*	*0.1% – 1.0%*	*0.9%*	*0.375%*
Quarantine fraction	*Percentage*	*0% – 100%*	*70%*	*52.5%*

We observe the calibrated value for the incubation period to be a little lower than the initial incubation period, yet still in the range of the RIVM. The calibrated infectious period is three days lower than the initial input value. The number of infected cases on 2 February is higher than initially assumed, just like the basic reproduction rate. The calibrated IFR is lower than the initial input value, but still in the range of Anderson, Heesterbeek, Klinkenberg & Hollingsworth (2020). The fraction of confirmed cases effectively quarantined is lower than initially assumed. We think the reason for this is that infected cases who are in quarantine have a hard time to not infect others in the same house. This assumption can be confirmed by the fact that approximately 50% of reported infection places is at home (Table 5).

Calibration values of indicators in the system dynamics model

In Table 15, we provide calibrated and actual values of indicators. Actual values were gathered on 20 January and may have changed a little thereafter. We observe that calibrated values for all indicators in the table come close to the actual values on 31 November.

Table 15 Calibrated values and actual values of indicators

	Confirmed	**Deaths**	**Hospitalizations**	**IC occupations**
Actual values	529,304	9,653	27,738	5,551
Calibrated values	529,719	9,583	27,548	5,510

In Appendix J, we provide the development of all indicators after calibration. We observe that the number of confirmed cases per day is 1,191 at the highest point of the first peak and 10,453 at the highest point of the second peak. These values come close to actual values. For the number of hospitalized cases, the number of IC occupations, and the number of deaths, the first peak is a little lower than the second peak. Whereas in Section 2.2, we observe the first peak of these indicators to be once as high as the second peak. This could indicate that the IFR and the hospitalization rate in reality changed over time. Further discussion of the development of these indicators will be done in Section 5.1, where we validate our system dynamics model.

Considering the calibrated values of our model are quite aligned with actual values, the actual total number of infected cases seems to be highly underestimated. In Section 4.3.4, we estimate the number of infected cases with help of the number of infectious cases. There we observe that the number of infectious cases per day during the first and the second peak is similar (see Figure 25). Consequently we believe the number of infected cases in the first and second peak are similar as well. The total number of infected cases on 31 November is 1,814,638 based on data in this figure, whereas the calibrated total number of infected cases in our model is much higher (3,006,600).

4.5 Conclusion

Now we answer research question 3: "How to include the key factors and measures in the model? "

In this chapter, we develop the system dynamics model that will be used to simulate policies in the next chapter. As fixed input parameters we use the incubation period, infectious period, basic reproduction rate, fatality ratio, initial number of infected cases, and the quarantine fraction. These factors are required to mimic spread, and are eventually calibrated in Section 4.4. Whether other factors (temperature, humidity, wind speed, adoption of government measures, and testing capacity) add predictive value for our model is determined with the backward regression method, since it is not clear yet whether these factors affect spread considerably. Factors and measures that add predictive value are included in a regression model that predicts the effective reproduction rate. The factors that remain after backward regression are temperature and adoption of government measures (staying home behaviour and traveling behaviour). All measures (school openings, caterings services openings, event allowance, and facemasks) are included in the model, because we want to express the effect of measures on spread. In the statistical analysis, we compare a linear, polynomial, and interaction model, to see which one provides the most accurate and reliable prediction. The first model including only parameters (factors and measures) in linear form, the second considering non-linear parameters as well, and the last considering relations between parameters by adding interaction terms to the linear model. The polynomial and the interaction model show better quality than the linear model with a higher adjusted R^2 and corrected R^2, and the interaction model also shows the lowest corrected MSE (test error) of all models. Yet both the polynomial and interaction model are considered to be overfit, and are thus not used in our system dynamics model. This remains $R_e{}^{linear}(t)$ to predict the effective reproduction rate and consequently the number of infected cases on a day. We predict the number of confirmed cases on day with the regression model $confirmed\ cases\ (t)$, with help of the number of tests and the number of infected cases on a day.

With all input above we develop the system dynamics model in Section 4.4, according to the SEIR model. Calibrated input values of parameters are provided in Table 14, and calibrated values of indicators in Table 15. We observe that obtained values for the number of confirmed cases, deaths, hospitalized cases, and IC occupations in our model are close to actual values on 31 November. Remarkably, the total number of infected cases is much higher than estimated in Section 4.3.4. Whereas we estimate 1,814,638 infected cases, the calibrated number of infected cases is 3,006,600.

5. Results

In this chapter, we simulate multiple policies to prevent spread of the virus. In Section 5.1, we first validate our system dynamics model. In Section 5.2, we outline the policies that are considered and provide necessary information about the setup of these policies in the system dynamics model. We provide the results of all policies in Section 5.3, and in Section 5.4 we perform a sensitivity analysis to test the robustness of our model.

5.1 Validation of the model

Before we simulate spread for all policies, we validate our system dynamics model to make sure that outcomes are valid. We use input data from 1 December until 31 December to simulate spread with our model and we compare the outcomes of the effective reproduction rate, the infectious population, the number of confirmed cases, the number of hospitalized cases, the number of IC occupations, and the number of deaths to the actual spread in December. We compare the infectious population instead of the number of infected cases, because the RIVM provides estimates for the infectious population per day (see Figure 25).

5.1.1 Comparing development of the effective reproduction rate

In Figure 27, we visualize the development of ${R_e}^{linear}(t)$ from the system dynamics model, and the development of the lower bound, upper bound, and average ${R_e}^{calculated}(t)$. We observe that the development of ${R_e}^{linear}(t)$ is similar to the development of the average value of ${R_e}^{calculated}(t)$.

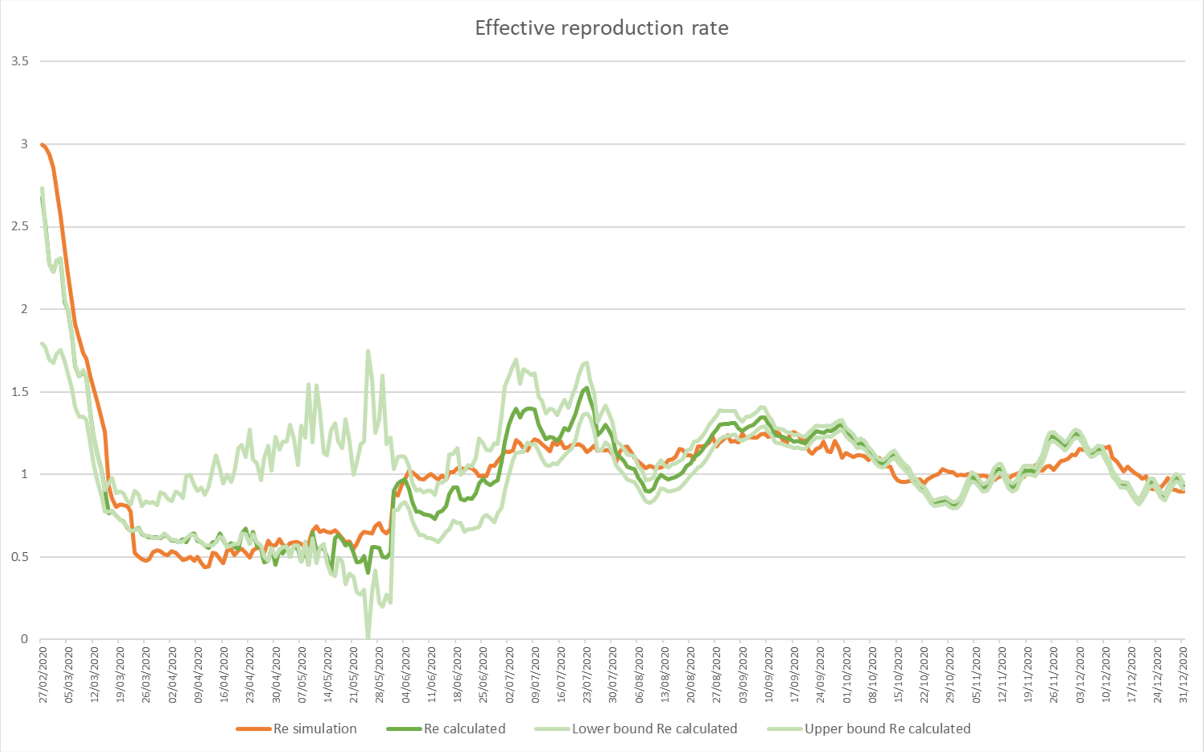


Figure 27 Effective reproduction rate from the system dynamics model and calculated effective reproduction rate

Remarkably, the highest observed value for ${R_e}^{calculated}(t)$ after implementation of the first government measures is 1.53, while the highest observed value for ${R_e}^{linear}(t)$ is 1.26. The minimum value is 0.40 for ${R_e}^{calculated}(t)$ and 0.44 for ${R_e}^{linear}(t)$. Further we observe that ${R_e}^{linear}(t)$ sometimes exceeds the lower or upper bound of ${R_e}^{calculated}(t)$, especially towards the end when the confidence interval becomes smaller. These observations indicate that ${R_e}^{linear}(t)$ has a hard time to provide outliers. The development of this regression model shows less peaks than $R_e^{\,calculated}(t)$. We think these differences might be caused by the input values that are used for our regression model $R_e^{\,linear}(t)$. In $R_e^{\,linear}(t)$, input values change less per day compared to input values of $R_e^{\,calculated}(t)$. For example, catering services might be closed for a long period, whereas the number of hospitalized or confirmed cases, the input values for $R_e^{\,calculated}(t)$, change every day. Unless these differences in fluctuations, we believe the approximations of the effective reproduction rate made with $R_e^{\,linear}(t)$ are valid and can be used for modelling the policies in Section 5.2.

5.1.2 Comparing development of the infectious population

In Figure 28, we compare the development of the infectious population from our system dynamics model with the development of the infectious population estimated by the RIVM (RIVM, 2020b). We observe that, before 1 June, the infectious population from our model lies between the lower and upper bound estimated by the RIVM. Besides, observed values from our model in the first and second peak are close to the average infectious population estimated by the RIVM.

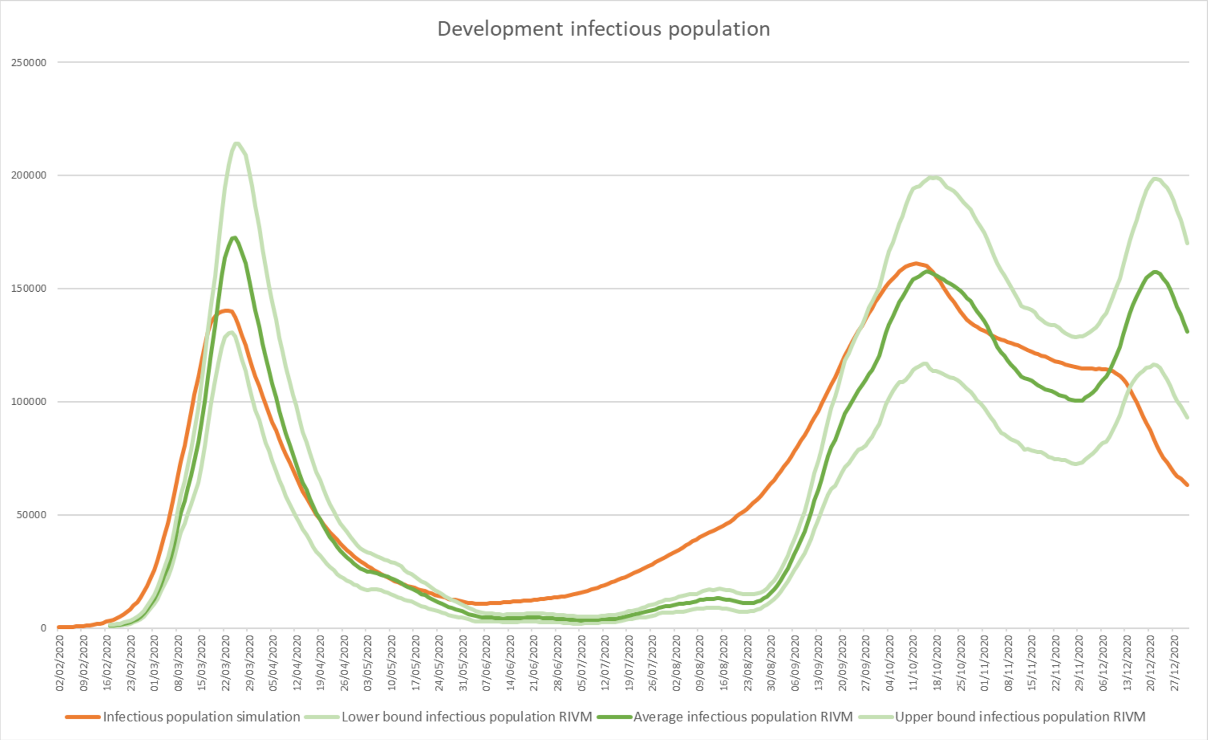


Figure 28 Infectious population from the system dynamics model and estimated infectious population by RIVM

From 1 June, we observe that the infectious population starts to increase earlier than the estimates of the RIVM. This indicates that the value of $R_e^{\,calculated}(t)$ we use after June might be a little too high. Since we use $R_e^{\,calculated}(t)$ based on the number of confirmed cases to train our regression model $R_e^{\,linear}(t)$ after 1 June, our calculation of $R_e^{\,calculated}(t)$ might deviate a little from actual $R_e(t)$. On 1 June, $R_e^{\,linear}(t)$ lies around the upper bound of $R_e^{\,calculated}(t)$. On this date, some measures were softened by the government. Unless the difference during the start of the second peak, the infectious period in our model at the highest point of the second peak is close to the estimate of the RIVM. After the second peak, we see a decline of the infectious population within the bounds until around 6 December. According to RIVM estimates, the size of the infectious population starts to increase again around 6 December and decrease just after the implementation of lockdown measures on 15 December. The infectious population from our simulation does not increase around 6 December but decreases from there.

Since $R_e(t)$ has a significant effect on the size of the infectious population, we think differences observed between $R_e^{calculated}(t)$ and $R_e^{linear}(t)$ in Section 5.1.1 cause the difference in observed and actual infectious population. When looking to $R_e^{calculated}(t)$ and $R_e^{linear}(t)$ in Figure 27, we see that the height of $R_e^{calculated}(t)$ around the end of November is a little higher than $R_e^{linear}(t)$, and declines to a somewhat lower point than $R_e^{linear}(t)$ in the middle of December.

5.1.3 Comparing development of the number of confirmed cases per day

In Figure 29, the development of the number of confirmed cases from our model is compared to the actual number of confirmed cases. The development of the actual number of confirmed cases per day is similar to the observed number of confirmed cases in our system dynamics model.

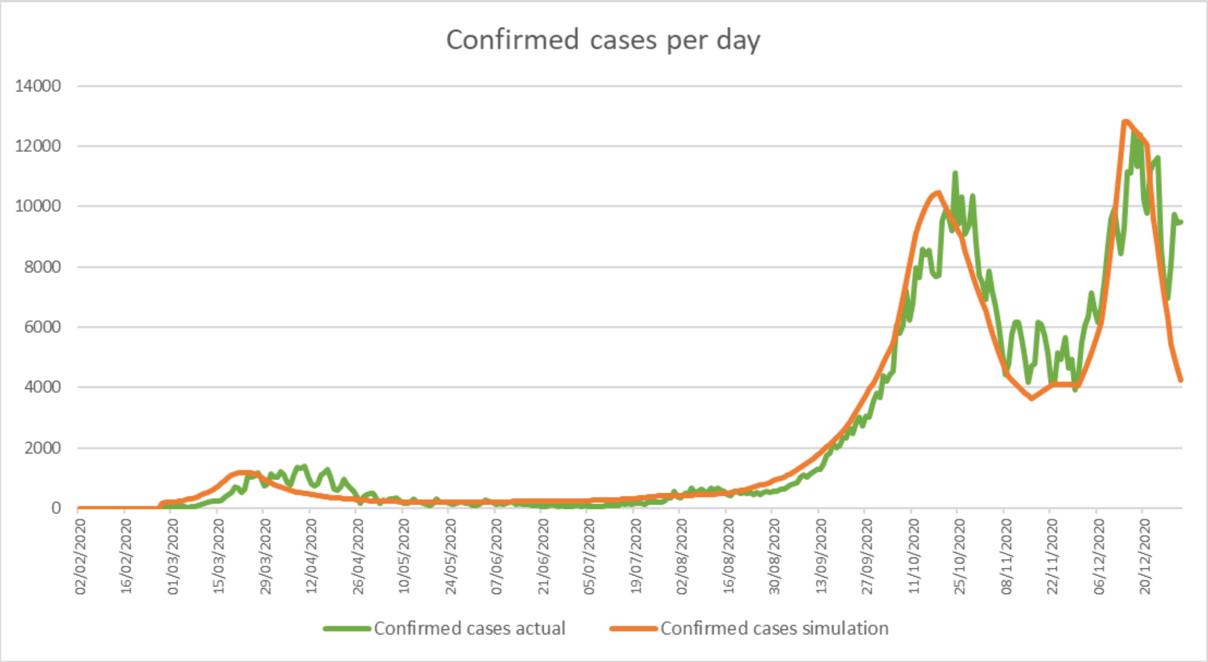


Figure 29 The number of confirmed cases from the system dynamics model and actual number of confirmed cases

We observe the first confirmed case on the same day, but the first peak occurs a little earlier in our system dynamics model. When we vary the incubation period to try to obtain a later peak, this does not lead to an earlier peak. For this reason, we believe the small difference in start of peaks is not caused by a deviation between the actual incubation period and the incubation period we use in our model, which seems an obvious cause. We think the actual cause to be a small difference between $R_e^{linear}(t)$ and the actual reproduction rate, because we observe in Section 5.1.2 that $R_e^{linear}(t)$ has a big impact on the infectious population and in Section 5.1.1 we observe that $R_e^{linear}(t)$ deviates a little from the average $R_e^{calculated}(t)$. Unless these small differences between the actual confirmed cases and the confirmed cases from our model, we believe our model can provide a valid estimate for the number of confirmed cases per day.

5.1.4 Comparing development of the number of hospitalized cases and IC occupations per day

In Figure 30 and 31, the development of the number of hospitalized cases per day and the number of IC occupations per day from our model are compared to the actual values. Whereas the observed number of confirmed cases per day in our system dynamics model does not deviate a lot from actual values, there exists a considerable difference between the development of the observed number of hospitalizations and IC occupations in our model and the actual numbers. We observe the first peaks of these indicators from our model are about once as low as actual values in the first peak.

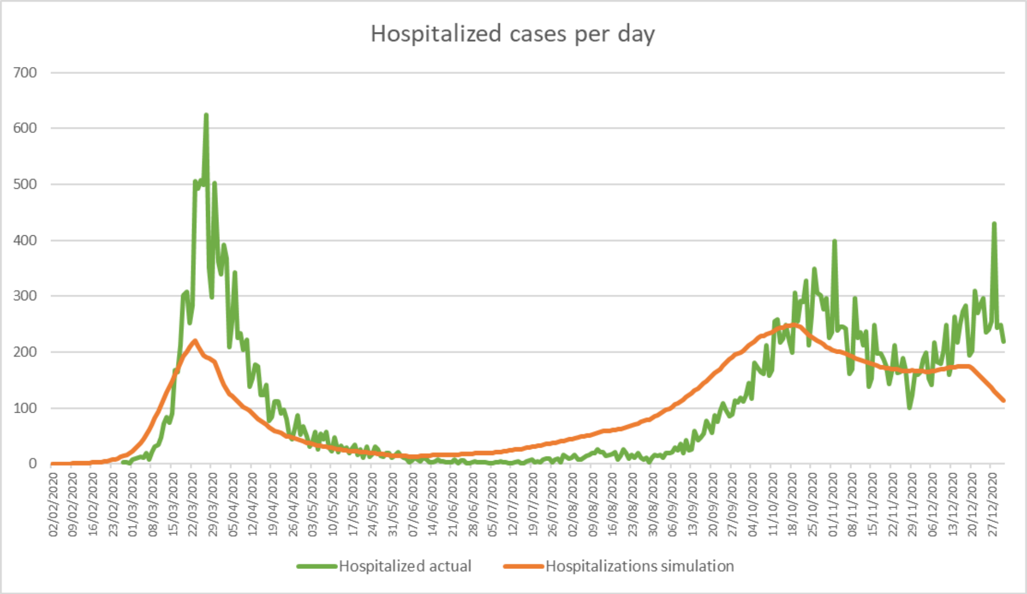


Figure 30 Number of hospitalized cases from the system dynamics model and actual number of hospitalized cases

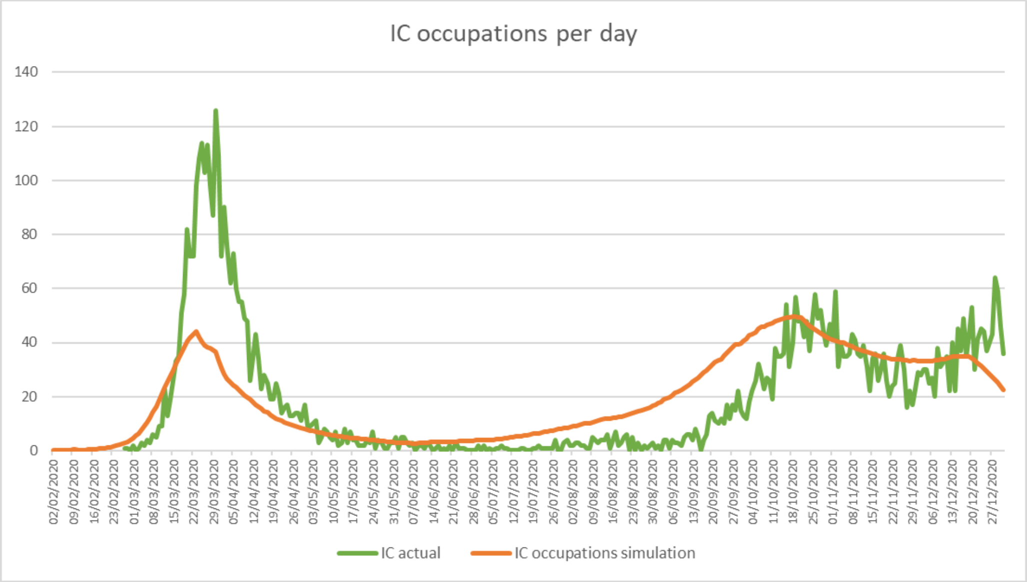


Figure 31 Number of IC occupations from the system dynamics model and actual number of IC occupations

The number of hospitalizations and IC occupations on a day are determined by the number of infected cases and the hospitalization rate. The number of infected cases follows from the infectious population. The height of the first and second peak of the infectious population are within the range of the RIVM (Section 5.1.2). For this reason, we believe that differences in actual and observed peaks of hospitalizations and IC occupations are not caused by a wrong estimate of the number of infected cases, but mainly by a change in hospitalization rate. In our system dynamics model we assume a fixed hospitalization rate of approximately 0.8% from the number of infected cases and 5.5% from the number of confirmed cases. We think the hospitalization rate might have considerably decreased between the first and second peak. Whereas the heights of the first and second peak of the infectious populations lie within the range of the RIVM, the second peak starts earlier than the actual peak. We observe the same for the development of the number of hospitalizations and the number of IC occupations per day. Therefore we believe a difference in the infectious population to be the main cause for an earlier start of the second peak for the number of hospitalizations and the number of IC occupations per day.

5.1.5 Comparing development of the number of number of deaths per day

In Figure 32, the development of the number of deaths per day from our model is compared to the actual development of the number of deaths. The development of the number of deaths per day from our model differs from the actual development. We observe the first peak to be around once as low as actual values in the first peak, similar to our observation for the number of hospitalizations and IC occupations per day.

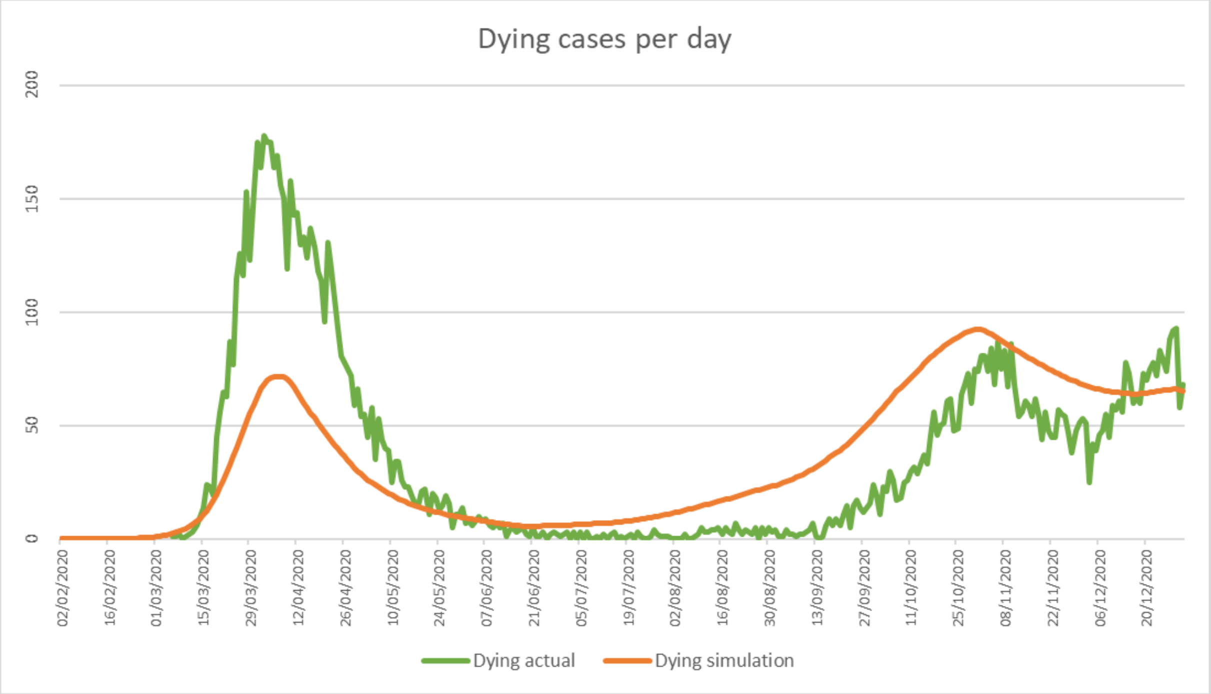


Figure 32 Number of deaths from the system dynamics model and actual number of deaths

The number of deaths per day follows from the number of infected cases and the IFR. Since we observe in Section 5.1.2 the size of the infectious population in the first and second peak to be quite accurate, we think the difference between the actual and simulated development of the number of deaths is caused by a decrease in IFR between the first and the second peak. Next to that, we believe the second peak of the number of deaths per day to start earlier due to an early increase of the infectious population, similar to the hospitalizations and IC occupations.

5.1.6 Comparing the outcomes of indicators

According to data from the RIVM and from association NICE, there are 808,906 confirmed cases, 11,627 deaths, 34,833 hospitalized cases, and 6,748 IC occupations on 31 December (NICE, 2020; RIVM, 2020f). For all indicators but the number of infected cases, values from our system dynamics model are somewhat lower than actual values (see Table 16). For the number of infected cases there is no data available.

Table 16 Validation and actual values of indicators on 31 December

	Infected	**Confirmed**	**Deaths**	**Hospitalizations**	**IC occupations**
Actual values	-	808,906	11,627	34,833	6,748
Validated values	3,437,300	790,411	11,614	32,557	6,511

We observe in Table 16 that observed values in our model are close to actual values for all indicators. The number of hospitalized cases deviates relatively the most from the actual value on 31 December, which we think is mainly caused by a change in hospitalization rate.

While actual values on 31 December come close to values from our model, the infectious population, number of deaths, number of hospitalizations and the number of IC occupations per day all show an earlier start of the second peak than observed in reality. For the number of deaths, hospitalizations and IC occupations we believe this earlier start compensates for the fact that values in the first peak is lower than the actual observed values in the first peak. Consequently, calibrated values on 31 November are quite accurate, but small differences between observed values and actual values on 31 December can be observed. Concluding, we believe the approximations made with our model are sufficient to be used for modelling the policies in Section 5.2, but we have to keep in mind the difference in development of the number of hospitalizations, IC occupations, and deaths when evaluating performance of policies.

5.2 Policies

According to Red Team, an independent group of experts that aims to prevent and fight COVID-19 in the Netherlands, there exist roughly four strategies to prevent a virus from spreading. These strategies are “Do nothing”, “Mitigation”, “Curbing”, and “Elimination”. Which one works best for public health, healthcare, and economy depends on the characteristics of the virus (RedTeam, 2020). In Section 5.3, we test the effect of proposed strategies of Red Team in the case of COVID-19. We do this by defining policies that determine the measures implemented in each strategy, and to what extent. Besides these policies, we test the effect of the measure facemasks and of testing capacity on spread. We shortly explain all policies and the policy that was chosen by the Dutch government in Section 5.2.1. In Section 5.2.2 and Section 5.2.3, we provide necessary information about how we model the policies.

5.2.1 The policies to prevent spread

As we mention above, we consider the following six policies:

1) Do nothing (“Niets doen”): Let the virus move freely.
2) Mitigation (“Mitigatie”): Accept circulation of the virus to a certain extent.
3) Curbing (“Indammen”): Strive for as little infections as possible, pursue every infection.
4) Elimination (“Uitroeien”): Make the virus disappear.
5) Facemasks
6) Testing capacity

The first policy, doing nothing, is not realistic for COVID-19. When the government would do nothing to prevent spread, we would have experienced major damage to public health, healthcare, and economy. In the second policy, mitigation, the main goal is to develop herd immunity. However, this strategy is stubborn in practice since it is not (yet) clear how many people develop immunity and for how long immunity lasts. With the third policy, curbing, it is acceptable to have an outbreak now and then, but it is not accepted to let the outbreak become uncontrollable. This means that infected individuals will be identified by quick testing and in-depth source- and contact investigation. This policy requires strict signal values and sufficient testing capacity. In some countries (e.g. China, Japan) this strategy has been proven successful. Yet in Europe this might require an approach where countries work together due to the high contact rate between European countries. The fourth policy eliminates the virus. The feasibility of this approach for COVID-19 is low in a short amount of time, since it requires world-wide cooperation. Next to the strategies proposed by Red Team, we test whether the implementation of facemasks is valuable to prevent spread, and whether the number of tests on day has effect on spread.

The policy that was chosen by the Dutch government is called “maximal control” (“maximaal controleren”). This policy can be seen as a mix of policies 2 and 3. The main focus of the government is to protect people who are vulnerable and prevent healthcare from overloading (RedTeam, 2020).

5.2.2 Modelling the policies to deal with spread of the virus

Here we explain how we model the proposed policies with our system dynamics model. In each policy, we start implementing measures on 15 March just like in reality.

Policy 1: Doing nothing

With this policy, no actions are taken to prevent spread and values for all input parameters remain 1 the entire period. This means that no facemask measures were implemented. We use low testing capacity in this policy. Values used for low testing capacity are provided per date in Appendix K.

Policy 2: Mitigation

With this policy, we use the number of IC occupations and hospital admissions on a day as a guideline to identify the measures to implement. Since the Dutch government uses hospital admissions and overload of IC occupations as a guideline to identify measures, we think this policy comes closest to the policy that is applied in the Netherlands.

In this policy, we determine input values of the measures school openings, catering services openings, event allowance, and the factors staying home behaviour and traveling behaviour based on the signal values in Table 17, and we use input values for these measures from Table 16. How these values exactly will be used is explained in Section 5.2.3. For facemasks and testing capacity we assume the same values as were actually implemented. In Appendix L, we provide the input values of this policy used in the system dynamics model per date. As can be seen, we start implementing lockdown measures on 15 March just like in reality. We do this to be able to better compare this policy to the actual policy.

Policy 3: Curbing

The Red Team denotes that the lesson learnt from countries who have successfully implemented curbing as policy use the number of confirmed cases as a guideline. The advantage of the number of confirmed cases as indicator to determine the policy is that this indicator indicates a change or trend in spread earlier than for example the number of hospitalizations. The drawback is that it takes a lot of time, money, and energy to successfully implement, due to the high testing capacity and strict follow-up of infected cases that is required. For this reason, we use the number of confirmed cases per day as indicator for when to implement softer or harder measures in this policy.

Similar to the mitigation policy, we determine parameter input values of the measures school openings, catering services openings, event allowance, and the factors staying home behaviour and traveling behaviour based on signal values in Table 18, and we use input values from Table 17. For curbing this will also be explained in more depth in Section 5.2.3. Since curbing requires a high testing capacity to quickly identify infected cases, we simulate this policy with a high number of tests (specification of number of tests per date in Appendix K). Besides, we use the same values for the measure facemasks as actually implemented. In Appendix L, we provide the input values of this policy for the system dynamics model per date. Similar to the mitigation policy, we start implementing lockdown measures on 15 March just like in reality.

Policy 4: Elimination

With this policy, we implement the strictest input values for all measures, meaning input values of all measures are 5. We assume that wearing facemasks is strongly urged in the first month (input value 4), from 12 March to 12 April, and mandatory thereafter (input value 5). Testing capacity is set high, to help quick tracking of infected cases. In practice this would mean that lockdown measures are implemented in the entire period.

Policy 5: Wearing facemasks

With this policy, we test what happens when wearing facemasks is mandatory in the entire period and when facemasks are not implemented at all. For the first, we assume a strong urge for wearing facemasks in the first month and mandatory wearing of facemasks thereafter. We do this because in reality it took some time to make wearing facemasks mandatory, due to government regulations. The values of all other parameters remain the same as in reality.

Policy 6: Testing capacity

With this policy, we test the effect of low testing capacity and of high testing capacity on spread. With a low testing capacity, the number of tests per week remains low in the whole simulated period. With a high testing capacity, the number of tests per week quickly becomes high. Input values of all other parameters remain the same as in reality. In Appendix L, development of values for the number of tests is provided per week. Here we distinguish actual testing capacity, low testing capacity, and high testing capacity.

An important parameter that influences the outcome of this policy is the fraction of confirmed cases that enters quarantine. A higher quarantine fraction means more confirmed cases in self-quarantine. We keep the quarantine fraction the same when modelling this policy. In the sensitivity analysis we consider changes in the quarantine fraction.

5.2.3 Input values for mitigation and curbing policy

In the previous section we mention that mitigation acts according to signal values of the number of hospital admission and IC occupations and curbing according to signal values of the number of confirmed cases. In this section we explain how we model these policies in the system dynamics model.

Parameter values per risk level

In order to model the mitigation and curbing policy proposed by Red Team, we use the route map ("routekaart") of the Dutch government (Rijksoverheid, 2020g). The Dutch government uses this route map to determine what measures to apply. This route map consists of five risk levels. These risk levels indicate the measures that are implemented by the government when spread of the virus is in a certain state. The five risk levels are alert ("Waakzaam"), alarming ("Zorgelijk"), serious ("Ernstig"), very serious ("Zeer ernstig"), and lockdown. In Table 17, we establish the values each measure should have per risk level. This is in accordance with the values per risk level in the route map of the Dutch government and the parameter values in Table 6.

Table 17 Parameter values of measures per risk level

	Before first measures	**Alert**	**Alarming**	**Serious**	**Very serious**	**Lockdown**
School openings	1	2	2	3	4	5
Catering service openings	1	2	3	4	5	5
Event allowance	1	1	2	3	4	5
Traveling behaviour	1	0.5	0.5	0.3	0.3	0.2
Staying home behaviour	0	4.5%	4.5%	9%	13.5%	18%

Values for traveling behaviour and staying home behaviour are not in the route map. The values we establish for these parameters are approximations, based on observed values from the past. The average values of staying home behaviour range from 0 to 18%, as we observe in Figure 19. Since we do not observe a clear pattern in the past for the serious, very serious and lockdown risk level, we assume a linear increase in staying home behaviour when going one risk level to the right. This way, the percentage of people staying home increases with 4.5% when we apply stricter measures (one risk level to the right) and decrease with 4.5% when we apply softer measures (one risk level to the left). We do observe similar values in the past for the alert and alarming risk level. For this reason, we assume the same percentage of people staying home in the alert and the alarming level.

For traveling behaviour we observe a pattern in the past (see Figure 20). Before any measures were implemented, traveling behaviour is 1 (baseline). During the first and second lockdown, traveling behaviour approaches 0.2. In other risk levels, values of traveling behaviour vary between 0.5 and 0.3.

Indicator signal values per risk level

Besides the measures that have to be implemented in each risk level, the values that indicators are allowed to take in each risk level are determined by the Dutch government. These so-called "signal values" are provided in Table 18. We assume that measures from a lower or higher risk level can only be implemented when measures of the current risk level are implemented for at least two weeks. We do this, because effects of adjusted measures are not immediately visible.

Table 18 Signal values of indicators per risk level

	Alert	**Alarming**	**Serious**	**Very serious**	**Lockdown**
Number of confirmed cases	< 1,218 p/d (< 7 per 100,000 residents)	1,218 – 3,654 (7 – 21 per 100,000 residents)	3,654 – 6,264 p/d (21 – 36 per 100,000 residents)	6,264 – 8,700 p/d (> 36 per 100,000 residents)	> 8,700 p/d (> 50 per 100,000 residents)
Reproduction rate	< 1	≈ 1	> 1	> 1	n/a.
Hospital admissions	< 40 (p/d)	< 40 (p/d)	40 – 80 (p/d)	> 80 (p/d)	> 160 (p/d)
IC occupations	< 10 (p/d)	< 10 (p/d)	10 – 20 (p/d)	> 20 (p/d)	> 40 (p/d)

For all risk levels but lockdown, value ranges in Table 18 are directly gathered from the route map of the government (Rijksoverheid, 2020g). For risk level lockdown we estimate signal values, because the route map does not provide these values. We do this based on signal values in other risk levels and observed values from the past, which we describe below.

For the number of confirmed cases, the difference in number of cases per 100,000 residents is mostly around 14 or 15 cases when going one risk level to the left or right. Besides, the actual number of confirmed cases per day was between 8,000 and 9,000 just before the implementation of a second lockdown on 15 December. The number of confirmed cases before the first (intelligent) lockdown was very low due to a low testing capacity, and is therefore not considered to provide a good estimate. For these reasons, we assume that lockdown measures apply when the number of confirmed cases is above 8,700 (50 cases per 100,000 residents).

On 16 March and 12 December, the number of hospitalizations was around 160. There is not a clear value for the number of IC occupations just before implementation of lockdown measures. In the route map, signal values for number of hospitalizations and IC occupations seem to be doubled when going one risk level to the right (stricter measures). For these reasons, we assume lockdown measures are implemented when the number of hospitalizations and IC occupations is once as high as in the very serious risk level. This means that lockdown measures apply when the number of hospital admissions is above 160 and the number of IC occupations is above 40 per day.

Modelling the mitigation policy

Above we explain how we use signal values to determine values of input parameters. According to these signal values, measures of the alert or alarming risk level apply when the number of hospital admissions on a day is below 40 and the number of IC occupations is below 10. Because the signal values for these two risk levels are the same, we assume that the alert risk level can only be reached when alarming measures have been implemented for at least one month. We act according to the serious risk level when the number of hospital admissions is between 40 and 80 and the number of IC occupations is between 10 and 20, the very serious risk level when the number of hospital admission is above 80 and the number of IC occupations is above 20, and the lockdown risk level when the number of hospital admission is above 160 and the number of IC occupations is above 40 per day.

Modelling the curbing policy

To determine the performance of curbing, we distinguish three ways to implement the curbing policy, which we refer to as curbing type 1, curbing type 2 and curbing type 3. The first type uses the signal values for confirmed cases in Table 18 to determine when to implement certain measures. The second type uses the signal values in Table 19. This means that when the number of confirmed cases increases above 1,218 per day, we apply lockdown measures until the number of cases declines below 1,218 again. When the number of confirmed cases per day is below 1,218, we apply the measures of the alert risk level. In the third type we only implement short and hard lockdowns of two weeks with two weeks of measures from the alert level thereafter.

Table 19 Adjusted signal values per risk level for curbing type 2

	Alert	**Alarming**	**Serious**	**Very serious**	**Lockdown**
Number of confirmed cases	< 1,218 p/d (< 7 per 100,000 residents)				> 1,218 p/d (> 7 per 100,000 residents)

5.3 Performance of the policies to prevent spread of COVID-19

In this section, we provide and evaluate results of all proposed policies from Section 5.2.1. We evaluate performance of the policies based on the number of infected cases and the number of days with strict measures, which we do in Section 5.3.2. We do not discuss results of the indicators deaths, hospitalizations, and IC occupations, since during validation we found that the development of these indicators deviates a little from actual development. Besides, these are fractions of the number of infected cases and thus the relative performance of these indicators is similar to the relative performance of infected cases. The indicator number of confirmed cases provides a wrong indication of spread when testing capacity is low and is therefore not used to indicate spread. This indicator can be used in the discussion, since in reality this indicator is used to express the number of infections.

5.3.1 Performance of policies

In Table 20, we provide the value of all indicators on 31 November per policy. We refer to the "actual policy" as the policy implemented by the Dutch government. In Table 21, we provide the number of days with strict measures per policy. Strict measures are considered to be measures of the serious, very serious, or lockdown risk level. We cannot provide a clear specification of the days with strict measures for the actual policy, facemasks, and testing capacity, because the Dutch government did not determine the policy exactly according to the route map in reality.

Table 20 Performance of indicators on 31 November per policy

Policy		**Total infected**	**Total confirmed**	**Total deaths**	**Total hospitalizations**	**Total IC occupations**
Policy 0	*Actual policy*	3,006,600	529,719	9,583	27,548	5,510
Policy 1	*Doing nothing*	10,989,500	5,869,190	40,619	106,962	21,392
Policy 2	*Mitigation*	2,734,690	265,183	9,398	25,839	5,168
Policy 3	*Curbing type 1*	4,417,520	1,016,010	15,502	41,952	8,390
	Curbing type 2	1,630,540	255,426	6,044	15,878	3,176
	Curbing type 3	2,436,290	444,190	9,098	23,817	4,763
Policy 4	*Elimination*	597,688	46,254	2,244	5,856	1,171
Policy 5	*Wearing facemasks ever*	7,607,750	2,644,600	28,001	74,007	14,801
	Wearing facemasks never	1,640,820	198,467	5,795	15,713	3,143
Policy 6	*Testing capacity high*	880,337	134,932	3,300	8,620	1,724
	Testing capacity low	3,694,050	311,810	11,102	33,074	6,615

Table 21 Number of days with strict measures per policy

Policy	**Days with strict measures**
Actual policy	+/- 200
Doing nothing	0
Mitigation	260
Curbing type 1	157
Curbing type 2	96
Curbing type 3	140
Elimination	264
Facemasks	+/- 200
Testing capacity	+/- 200

Next to total values of indicators on 31 November, we graphically show development of the total number of infected cases and the number of confirmed cases per day from 2 February until 31 November (in Figure 33 and 34), to get better insight in the relative difference in performance of policies. We observe in Figure 34 that the number of confirmed cases for doing nothing is very high in the first peak, causing development of other policies to be somewhat less clearly visible. The remaining policies show similar heights of peaks, but on different moments in time.

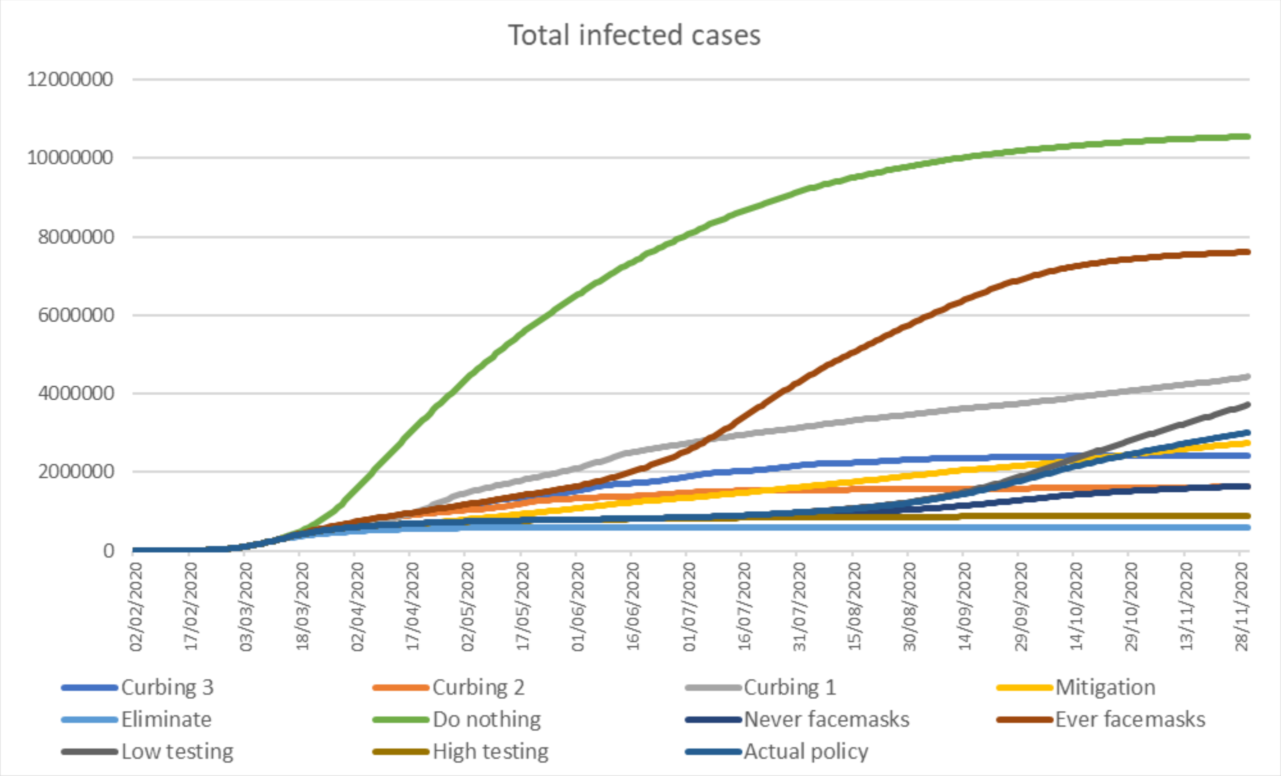


Figure 33 Total infected cases of all policies

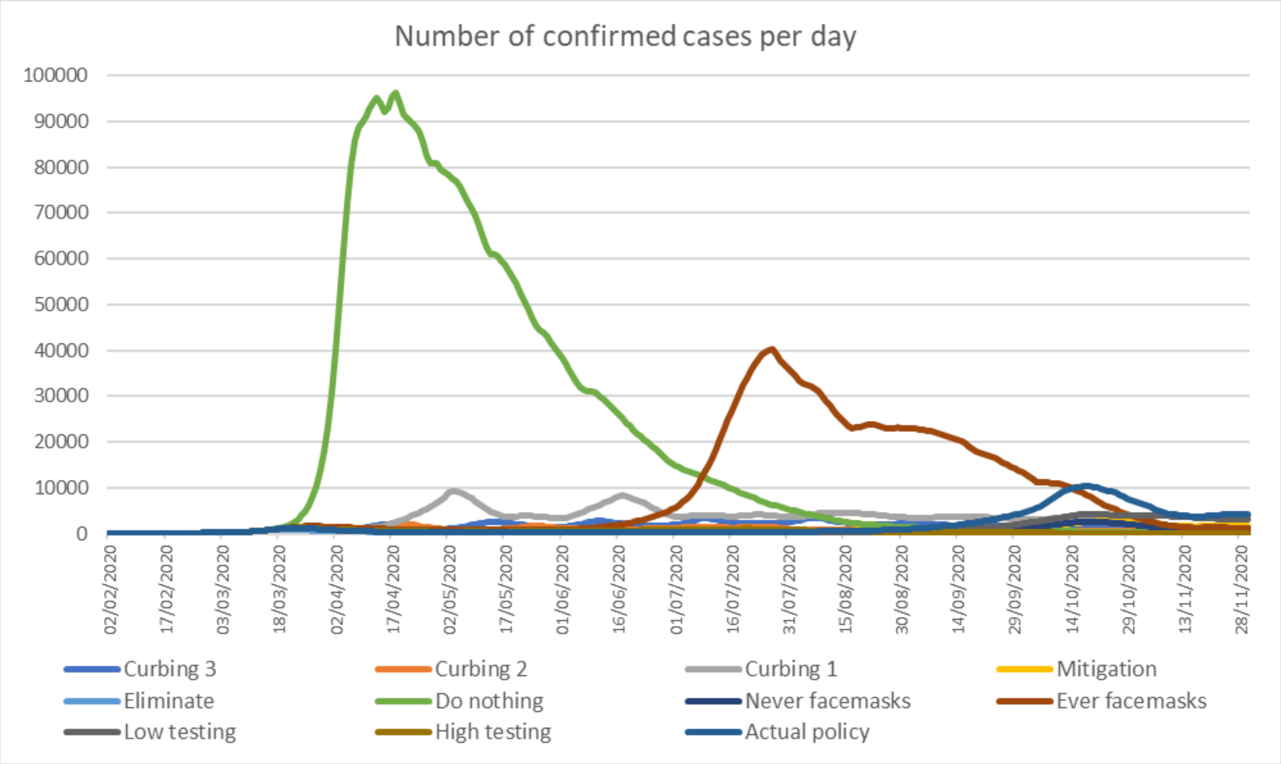


Figure 34 Number of confirmed cases per day of all policies

5.3.2 Conclusion of performance of policies

Below we shortly discuss the results of the non-feasible and feasible policies to prevent spread of COVID-19, to obtain a better understanding in the performance of all policies. Non-feasible policies are considered to be elimination and do nothing.

Performance of non-feasible policies

As we expect, the policy elimination results in the lowest values of spread and the policy doing nothing results in the highest values of all policies. When the government would have done nothing to prevent spread, the total number of infected cases on 31 November would be almost 11,000,000 and the total number of deaths would be 40,619. When the government would have chosen to eliminate the virus, the total number of infected cases would be 597,688 and the total number of deaths 2,244. We assume that elimination and doing nothing are not feasible policies to prevent COVID-19 from spreading. Both policies would cause major damage to public health, healthcare, and economy. Elimination because all schools, catering service, and events would be closed for a long time, and doing nothing because hospital pressure would overflow.

Performance of feasible policies

We observe in Table 20 that when testing capacity would have been high from the start of the pandemic, observed spread with the actual policy would have been greatly reduced. High testing capacity shows after elimination lowest values for spread of all policies, with curbing type 2 thereafter. The number of days with strict measures on the other hand is relatively high for the policy high testing. Curbing type 2 shows the lowest number of days with strict measures of all feasible policies.

The policy with high testing capacity results, after elimination, in the lowest total number of infections (880,337). This is more than 2,000,000 less than the number of infected cases with the actual policy. The policy with low testing capacity results, after doing nothing and curbing type 1, in the highest total number of infected cases (3,694,050), which is approximately 700,000 more than the actual number of infected cases. Whereas the number of infected cases with high testing and low testing differ greatly, the total number of confirmed cases for the policy with high testing capacity is just a little lower than the policy with a low testing capacity (134,932 versus 311,810). The number of days with strict measures of both policies is equal to the number of days with strict measures in the actual policy.

Remarkably, the policies that use indicator values from the route map of the government to determine the policy, mitigation and curbing type 1, show the highest number of days with strict measures (of feasible policies). The policy mitigation results in a higher number of days with strict measures than all three types of the curbing. This could indicate that the signal values from the route map are not sufficient to efficiently prevent spread of the virus. Curbing type 2 and 3 apply only lockdown measures and have significantly less days with strict measures. At the same time, the values of all indicators for curbing type 2 and curbing type 3 are lower than the values of indicators from the policy mitigation. Consequently, the number of hospitalizations, IC occupations, and deaths observed with these policies is low. Curbing type 3 has more days with lockdown measures and results in somewhat higher values for indicators than curbing type 2, but performs better than mitigation type 2 and the actual policy.

According to the linear regression model, the measure facemasks has a lot of impact on the spread of the virus. When facemasks would have been implemented from the start of the pandemic, the total number of infected cases would have been more than once as high than with the actual policy. When facemasks would not have been implemented at all, the number of infected cases would have been almost half of the number of cases from the actual policy. We believe that these results are remarkable. In the discussion we discuss this observation in more depth.

5.4 Sensitivity analysis

In this section, we test the robustness of our model by making changes in values of sensitive variables. Sensitive variables are parameters that can significantly change the outcome of the model. We test to what extent this is the case for the basic reproduction rate, initial infected cases, quarantine fraction, and the fraction of asymptomatic cases. These parameters are considered sensitive, because values of these parameters vary a lot in literature. We determine robustness of our model with help of Monte Carlo simulation methods. Next to robustness of our model, we test robustness of the results from Section 5.3 by making changes in the quarantine fraction. We consider only this parameter because it is the only sensitive input parameter that can be influenced in reality. Other (calibrated) input parameters (incubation period, infectious period, basic reproduction rate, initial infected and the infection fatality ratio) cannot be influenced.

5.4.1 Monte Carlo simulation

Monte Carlo simulation methods are a way of computing results that rely on repeated random sampling. These methods are especially useful when studying systems that have a high number of degrees of freedom. They are able to provide a routinely better prediction than human intuition. Monte Carlo is a good method to apply on our model, since our model relies on many uncertain inputs (Vensim, n.d.). The program Vensim provides multiple options for setting up a sensitivity analysis. Two of these options are univariate sensitivity and multivariate sensitivity. With univariate sensitivity analysis we are able to study the effect of changing one parameter's input value. To determine input values, the program automatically samples a set of numbers from a bounded domain that can be determined on forehand. With multivariate sensitivity analysis we are able to study the effect of changing multiple parameter's input values at the same time. In Section 5.4.2, we apply multivariate analysis to see how sensitive the model is when we simultaneously change parameters for the actual policy. We apply univariate sensitivity analysis in Section 5.4.3 to get more insight in sensitivity of individual parameters. In both the univariate and multivariate sensitivity analysis, we run 500 simulations per changing parameter. In the multivariate sensitivity analysis this results in 1500 simulations in total. As output, Vensim graphically provides confidence intervals for indicator(s) of interest. From all simulations, 100% of the simulation are in the grey part, 95% in the blue part, 75% in the green part, and 50% in the yellow part that we observe in the sensitivity graphs. For each sensitivity test, we only provide the sensitivity graph of indicator(s) that are considered to be relevant.

5.4.2 Multivariate analysis

Here we change the basic reproduction rate, initial number of infected cases and quarantine fraction at the same time. Since our model is very sensitive, we let Vensim randomly pick numbers from a normally distributed domain. We do this because we want the program to know which values are likely to fit reality. In Table 22, we provide the distribution of the parameters that are considered in this multivariate sensitivity analysis. We make slight changes to the calibrated input values of the basic reproduction rate, initial infected cases, and quarantine fraction. Since we initially do not distinguish between asymptomatic and symptomatic cases in our model, this parameter is not taken into account in the multivariate analysis. We study what happens when we adjust our assumption of the fraction of asymptomatic cases with univariate analysis in Section 5.4.3.

Table 22 Distributions for parameters in multivariate analysis

Parameter	**Current input value**	**Distribution {min; max; mean; standard deviation}**
Basic reproduction rate	3	{2.5; 3.5; 3; 0.15}
Initial infected cases	600	{100; 1,100; 600; 100}
Quarantine fraction	52.5%	{0%; 100%; 50%; 10%}

In Appendix M, sensitivity results of the multivariate analysis can be found for all indicators. We observe that the simultaneous changes made in input values of the basic reproduction rate, initial infected cases, and the quarantine fraction have a major impact on the outcome of all indicators. According to the sensitivity graphs, spread in the Netherlands would have stopped after the first peak when the reproduction rate lies around 2.5, the initial number of infected cases is low, and the quarantine fraction approaches 100% (assuming it would not come back to the Netherlands another time). When the reproduction rate approaches 3.5, the initial number of infected cases is high, and the quarantine fraction approaches 0%, spread could become once as high as actual values. Results of this multivariate analysis are discussed in more depth in Appendix M.

5.4.3 Univariate sensitivity analysis

With help of univariate sensitivity analysis, we consider changes of input values for one parameter at the time. We do this for the basic reproduction rate, the number of infected cases, the quarantine fraction and the fraction of asymptomatic cases. Of these parameters, only the quarantine fraction can be influenced in reality. For the parameters that cannot be changed in reality we randomly pick values from a pre-specified domain. For the quarantine fraction we also use specific values to get insight in the effect of this parameter on the outcome of feasible policies.

Basic reproduction rate and initial infected cases

According to the RIVM, the basic reproduction rate ranges between 2 and 2.5. The basic reproduction rate in our model was calibrated to be 3. The initial number of infected cases was estimated to be 111 on 14 February, based on an IFR of 0.9% and the first confirmed death on 6 March, and calibrated to be 600. When we use a different basic reproduction rate and/or initial infected cases, impact on spread is large. We apply four sensitivity tests and compare their outcomes to get more insight in the effect of these parameters. Specification of these tests is provided in Appendix N. In the figures in Appendix N, we provide the number of confirmed cases per day from all four sensitivity tests.

We conclude from these sensitivity tests that when the reproduction rate lies between 2 and 2.5, the number of infected cases on 2 February has to be a lot higher than 600 to obtain actual spread. Since we think it would be strange that all these cases remained unidentified until 27 February, we believe that it is more likely to have a higher basic reproduction rate than a higher number of infected cases on 2 February. We believe that the combination of the calibrated input values that are used to simulate the policies in Section 5.2, a basic reproduction rate of 3 and 600 infected cases on 2 February, is a likely combination of input values.

Quarantine fraction

In Section 5.3, we observe that the number of tests per day has major effect on spread, based on the outcomes of the policy testing capacity. The number of tests per day affects the number of confirmed cases and consequently the number of cases entering self-quarantine. To analyze the effect of the fraction of confirmed cases that is effectively quarantined in more depth, we vary the quarantine fraction in several sensitivity tests. Specification of these tests is provided in Appendix O.

We see that when we vary the quarantine fraction between 0% and 100%, this has major impact on spread. The number of confirmed cases per day ranges from approximately 300 to 20,000 is the first peak and from zero to almost 68,000 in the second peak. The total infected cases ranges from approximately 500,000 to 7,500,000 and the total deaths ranges from approximately 2,000 to 26,000. We observe that all policies show the lowest total number of infected cases with a quarantine fraction of 80%. These sensitivity tests show that when the number of tests per day and the quarantine fraction are both high, relatively low spread and a high number of days with soft measures (alert or alarming risk level) are achieved at the same time. No matter the policy that is chosen.

Fraction of asymptomatic cases

We determine the impact of the fraction of asymptomatic cases by changing the infectious period. This means that when the fraction of asymptomatic cases would be higher than our assumption (40%), the average infectious period would become lower than 6 days. When the fraction of asymptomatic cases would below 40%, the average infectious period would become higher than 6 days.. In the univariate analysis, we consider two sensitivity tests. In the first we randomly pick values for the infectious duration from a normally distributed domain with mean 7.4 days, standard deviation of 1 day, a minimum of 4.4 days, and a maximum of 10.4 days. In the second test we consider a mean of 6 days, standard deviation of 0.2 days, a minimum 5 days and maximum 7 days. In Appendix P, we explain these choices in more detail, and we provide the sensitivity graphs for these two sensitivity tests.

Based on the results of these tests we conclude that the infectious period and thus the fraction of asymptomatic cases is a sensitive parameter in our model. Small changes significantly influence the outcome. We believe the considerable difference between duration of the infectious period according to our findings in literature and duration of the infectious period in our model can have two causes. It could mean that infected cases are on average more infectious around symptom onset than currently found in literature, or it could mean that fraction of the number of cases without symptoms or with soft symptoms is higher than currently assumed in literature.

5.5 Conclusion

Now we answer research question 4: “What policy can effectively reduce spread of the virus in the Netherlands?”

After validating the model in Section 5.1, we propose six policies to prevent spread in Section 5.2. These are doing nothing, mitigation, curbing, elimination, facemasks, and testing capacity. Doing nothing and elimination are not considered to be feasible for the Netherlands, since these policies would cause major to public, health, healthcare, and economy. Mitigation accepts circulation of the virus to a certain extent, while curbing strives for little infections as possible. We let mitigation determine the policy based on signal values for the number of hospital admissions and IC occupations, and curbing on signal values for the number of confirmed cases per day. The government developed a route map where these signal values are provided. We transform this route map by removing measures we do not consider (see Table 18). The measures implemented when certain signal values are reached are provided in Table 17. For curbing we consider two different versions additionally, that consider different signal values to determine the policy. These two versions, curbing type 2 and curbing type 3, only apply strict measures. Results of the system dynamics model show that curbing type 2 and curbing type 3 both outperform mitigation and curbing type 1 in terms of spread and number of days with strict measures. Curbing type 2 and curbing type 3 show the lowest number of days with strict measures of all policies. The policy with high testing capacity shows, after elimination, the lowest number of infected cases, while the policy with low testing capacity shows, after doing nothing and curbing type 1, the highest number of infected cases. The number of days with strict measures with the policy testing capacity is relatively high. Facemasks have a significant impact on spread in our model, yet we question this outcome. Results are discussed in more depth in the next chapter.

The sensitivity analysis shows that our model is very sensitive. Small changes in input values of parameters can have significant impact on spread. As an example, to mimic spread, the basic reproduction rate has to be higher than 2.5, or the initial infected cases on 2 February has to be higher than 1,000. The last is considered to be unlikely, and we thus assume a basic reproduction rate of 3 with 600 initial infected cases. Furthermore, we observe the quarantine fraction to have a significant impact on the performance of policies. This will be discussed in more depth in the next chapter.

6. Conclusions & Recommendations

In this chapter, we answer the main research question:

What are the impact of government measures on the spread of COVID-19 in the Netherlands and how can we learn from this for future outbreaks and pandemics?

We answer this question in Section 6.1. In Section 6.1.1, we answer how the outcome of our model can be used in the Netherlands. In Section 6.1.2, we discuss how our findings can be used in the fight against the spread of COVID-19 worldwide, and in Section 6.1.3 to what extent our outcome can be used in future pandemics. In Section 6.2, we compare our findings to findings in literature, in Section 6.3, limitations of our study are provided, and in Section 6.4, our recommendations from this study.

6.1 Conclusions

6.1.1 How can the outcome of the models we used in the Netherlands

Below we first explain how the outcome of our model $R_{e\ linear}(t)$ can be used. Thereafter we provide main findings of our proposed policies, to identify the best policy to prevent spread of COVID-19 in the Netherlands.

Outcome of the regression model

In our system dynamics model, we use the regression model $R_{e\ linear}(t)$ to predict virus transmission. Below we shortly describe the main findings from this regression model.

- The extent to which people adopt government measures has a major impact on the effective reproduction rate and consequently on spread. The normalized estimate for β_i of staying home behaviour is -0.663 and the normalized estimate for β_i of traveling behaviour is 0.477.
- Stricter measures for event allowance considerably reduce spread, with a normalized estimate for β_i of -0.434.
- Higher ambient temperature decreases spread, with a normalized estimate for β_i of -0.339.
- Stricter measures concerning school openings (normalized estimate for β_i = 0.265), catering services openings (normalized estimate for β_i = 0.085) and facemasks (normalized estimate for β_i = 0.286) increase spread in our model, which seems not to be in line with our findings in literature. We believe the reason for these positive relations is that the effect of a certain measure (e.g. school openings) is partially expressed with another variable (e.g. traveling behaviour). These interactions can be expressed with interaction terms, yet our interaction model is considered to be overfit and is thus not used (discussed in more depth in Section 6.2).
- Other parameters we study that are not included in this regression model do not show a statistically significant relation to the effective reproduction rate, with humidity a *p*-value of 0.311 and wind speed a *p*-value of 0.944. Therefore these parameters are considered to not add predictive value for the model.

Performance of policies

We propose six policies to prevent spread of the virus, doing nothing, mitigation, curbing, elimination, facemasks, and testing capacity. Of these policies, doing nothing and elimination are not considered to be feasible, since they would cause major damage to economy and social life. We assess the quality of the remaining policies based on spread, with the total number of infected cases on 31 November, and based on the number of days with strict measures, considered to be serious, very serious, or lockdown measures. Besides the number of infected cases, we provide the number of confirmed cases, because this indicator is used to express the number of infected cases by the government. Since this indicator can deviate a lot from the number of infected cases, the number of confirmed cases provides a distorted view on spread when testing capacity is low. All results are provided in Table 23 per policy.

Table 23 Summary of performance per policy

Policy		**Total infected**	**Total confirmed**	**Total days with strict measures**
Policy 0	*Actual policy*	3,006,600	529,719	+/- 200
Policy 1	*Doing nothing*	10,989,500	5,869,190	0
Policy 2	*Mitigation*	2,734,690	265,183	260
Policy 3	*Curbing type 1*	4,417,520	1,016,010	157
	Curbing type 2	1,630,540	255,426	96
	Curbing type 3	2,436,290	444,190	140
Policy 4	*Elimination*	597,688	46,254	264
Policy 5	*Wearing facemasks ever*	7,607,750	2,644,600	+/- 200
	Wearing facemasks never	1,640,820	198,467	+/- 200
Policy 6	*Testing capacity high*	880,337	134,932	+/- 200
	Testing capacity low	3,694,050	311,810	+/- 200

The goal of this study is to reduce the impact of preventive measures on social life and economy and to keep hospitals and IC units from overflowing, while effectively reducing spread at the same time. We conclude by comparing performance of all policies that implementation of strict measures is useful to reduce spread of COVID-19 effectively, especially when these measures are implemented early. Besides, we observe that a high testing capacity can really be valuable, but this requires the number of infected cases effectively entering self-quarantine to be high enough. We conclude this from our findings in the sensitivity analysis, where we compare the effect of a high testing capacity with a quarantine fraction of respectively 20%, 52.5% and 80%. The policy with high testing capacity performs very well in terms of spread, yet when the quarantine fraction becomes low, curbing type 2, curbing type 3 and mitigation outperform this policy. When the quarantine fraction becomes high, policies with actual, high, and low testing capacity all outperform remaining policies in terms of spread. Thus to effectively reduce spread, high testing capacity and high quarantine fraction together are even more important than the strictness and timing of implementation of measures. Yet the number of days with strict measures is high for these policies (see Appendix O). We therefore believe curbing type 2 to be the best policy to reduce spread. With this policy, strict measures are implemented early and a high testing capacity quickly identifies infected cases to effectively reduce spread, and this policy keeps the number of days with strict measures relatively low, which is better for economy and social life.

Remaining important findings from our study are:

- Since the number of days with strict measures with the policies mitigation and curbing type 1 are considerably higher than with curbing type 2 and curbing type 3, we believe the signal values in the route map of the government are not sufficient to reduce spread quickly.
- We observe the policy facemasks to provide remarkable values for the number of infected cases. The cause of this remarkable performance is its estimated effect in the regression model, which we believe is somewhat off. This will be discussed in more detail in Section 6.3.
- The calibrated number of infected cases on 31 November is significantly higher than the estimated number of infected cases on 31 November.

6.1.2 How can the outcome of the model be used worldwide

The regression models in this study are developed with data from spread in the Netherlands. This means that the estimated effects of key factors and measures on the spread of COVID-19 are specifically applicable for the Netherlands. We do not study to what extent estimated effects of these key factors and measures apply for other populations, while the effect of measures can have a drastically different impact when implemented in a different country (Haug et al., 2020). Differences in effects of parameters between countries can be caused by many factors. One of them being population density, which we found to have a considerable impact on spread. Another being the age-distribution of a population. For these reasons, we cannot substantiate how well our model predicts spread in other countries. Fortunately, when acquiring data from another population, estimates can easily be updated by applying multiple linear regression on new data. Next to the regression models, the system dynamics model can easily be adapted to mimic spread for another population.

Whereas the regression models and system dynamics model might have to be adapted before they can be used to mimic spread in another population, we believe that performance of our proposed policies can be used in other countries.

6.1.3 How can the outcome of the model be used in future outbreaks and pandemics

Just like our regression and system dynamics model can be easily adapted to mimic spread for another population, our models can be adapted for future outbreaks and pandemics of different viruses.

An example for a likely cause of a future outbreak in the Netherlands is a recent change in the pandemic of COVID-19, with new variants that have arisen in different parts of the world. The most well-known variant in the Netherlands is the VK-variant (VK = Dutch abbreviation for United Kingdom; “Verenigd Koninkrijk”), which thanks its name to the place where it was first recognized. This variant appeared in the Netherlands in the end of December. The VK-variant is more contagious than previous variants in the Netherlands, and according to a recent advice of the Outbreak Management Team (OMT), a team who involves identifying and advising on outbreaks and threats of infectious diseases, more often leads to symptoms (Rijksoverheid, 2020d). These changes require the basic reproduction rate and the average infectious period of this variant to have higher values than calibrated values in our model. At moment of writing, the VK-variant is the most common variant of COVID-19 in the Netherlands that could, even with stricter measures, lead to a future outbreak. Since these new variants could cause new difficulties in handling with spread of COVID-19, a different policy might be required. Yet we believe that testing capacity and self-quarantine remain important, if not more important. Next to new variants of the virus, the start of vaccination in the Netherlands on 8 January could cause a different policy to be more applicable to effectively reduce spread. We do not study these changes of COVID-19 in our research.

In addition to the adaptations we mention in Section 6.1.2, input values for disease factors (i.e. basic reproduction rate, initial infected cases, incubation period, infectious period, and fatality ratio) might require different input values when we have to deal with a pandemic of a different virus in the future.

6.2 Discussion

6.2.1 Outcome of the regression models

Findings of Haug et al. (2020) suggest that there exists no magical measure that is able to decrease the effective reproduction rate below one on its own. A suitable combination of measures is necessary to effectively prevent spread of the virus (Haug et al., 2020). With help of the backward regression method, we found the measures adoption of government measures, event allowance, school openings, and facemasks, and the factor temperature together to provide a quite accurate prediction of the effective reproduction rate. Due to our particular interest, we added the measure catering services openings to this model. In the study of Brauner et al. (2020), who use data from January until May 2020, measures that can considerably reduce spread of COVID-19 are school closure, limiting gatherings to 10 people or less (i.e. event allowance), and closing face-to-face businesses (e.g. catering services). Our findings for school openings, catering services openings, facemasks, and adoption of government measures differ somewhat from findings in literature, which we discuss below.

- **School openings:** In our regression model, stricter measures concerning school openings increase the effective reproduction rate. We did not find any study that has evidence for school openings to considerably increase or reduce spread on its own. Recent modelling studies of COVID-19 predict that closing schools alone prevents spread much less than other measures (Viner et al., 2020). The effect of closing schools even seems to have changed during the pandemic (Haug et al., 2020). According to Rozhnova et al. (2020), impact of this measure depends on remaining opportunities to reduce non-school-based contacts. If measures other than closing schools are undesired or already implemented and the effective reproduction rate is still close to 1, the additional benefit of closing schools may be considerable (Rozhnova et al., 2020). Since all studies found school closure to decrease the effective reproduction rate, which collides with our findings, we believe there exists an interaction between school openings and traveling behaviour. We think these two together lead to a negative effect on the effective reproduction rate. This assumption can be confirmed by the findings in our interaction model, in which we find a statistically significant interaction between the two (see Appendix G). Unfortunately, our interaction model does not provide a sufficient prediction for the effective reproduction rate and is thus not used in our study.
- **Catering services openings:** The effect of catering services openings on spread of COVID-19 is low according to $R_{e\ linear}(t)$. According to Brauner et al (2020), closing among others catering services has moderate to small effect on spread. Yet in their study, catering services are considered to have a high infection risk. We think the reason for our findings to deviate from the findings of Brauner et al. (2020) could be caused by the fact that our regression model does not have sufficient data to determine the exact effect of closing catering services. This means that, similar to the measure school openings, the measure catering services openings might express its effect to a certain extent with help of staying home behaviour and/or traveling behaviour. In our interaction model there is a statistically significant relation between catering services openings and traveling behaviour, which supports this statement.
- **Facemasks:** Implementation of facemasks makes the effective reproduction rate increase significantly according to $R_{e\ linear}(t)$. We think the cause for this might be that people perceive a significant decrease in preferred distance from another person when this other person wears a facemask, in comparison to a person who does not wear a facemask. This assumption is supported by several studies (NU.nl, 2020g)(Cartaud, Quesque, & Coello, 2020). However, there does not exist clear evidence for facemasks to increase spread either, especially not to the extent in our regression model (NU.nl, 2020g). Concluding, there remain many mysteries about the working of facemasks still (Tabatabaeizadeh, 2021).

- **Adoption of government measures:** Our model expresses adoption of government measures with the parameters staying home behaviour and traveling behaviour. We find that the effect of people stay at home has a relatively large impact on the effective reproduction rate. We believe the reason for this is that the regression model partially expresses the effect of government measures (school openings, catering services openings, event allowance, facemasks) with the parameter staying home behaviour, because in reality, more people stay at home when government measures become stricter. We assume the same for the parameter traveling behaviour. However, traveling behaviour only takes into account the amount of traveling with public transport and no other ways of movement, leading to a somewhat smaller effect on the effective reproduction rate compared to staying home behaviour (normalized estimate $\beta_{staying\ home\ behaviour}$ = -0.663, normalized estimate $\beta_{traveling\ behaviour}$ = 0.477). The above findings explain the reason for the measures school openings, catering openings, and facemasks to be positively related to the effective reproduction rate, while findings in literature propose the opposite. In our study, all parameters in $R_{e\ linear}(t)$ are adjusted at the same time, and thus the effect of a certain parameter (e.g. school openings) that is partially expressed with another parameter (e.g. traveling behaviour) neutralize each other. Despite these drawbacks in our linear model, we believe that $R_{e\ linear}(t)$ is able to provide a valuable estimate of the effective reproduction rate for the purpose of our study.

6.2.2 Outcome of the system dynamics model

Calibrated settings of input parameters of the system dynamics model

From all calibrated parameters in our model, we believe the basic reproduction rate and initial number of infected cases to be the only parameters that need to be discussed. Findings for the basic reproduction rate in our model, calibrated to be 3, deviate from findings of the RIVM, who estimate the basic reproduction rate to be between 2 and 2.5. In literature, estimates for the basic reproduction rate of COVID-19 vary a lot, especially between countries. In one study, the basic reproduction rate is estimated to be 3.15 (He et al., 2020). In another study, the maximum basic reproduction rate in the Netherlands is even estimated to be approximately 6 (Linka, Peirlinck, & Kuhl, 2020). Furthermore, while the first cases of COVID-19 in the Netherlands were confirmed on 27 February, there exists a possibility that the first cases were already infected at the end of January (RTLnieuws, 2020a). Yet exact numbers are hard to discover. When the basic reproduction rate in our model is lower, this would lead to a higher initial number of infected cases. We cannot substantiate the initial number of infected cases to be higher, since we believe it would be rare that more than 600 cases remain unidentified until 27 February. For this reason, we believe a basic reproduction rate of 3 with 600 infected cases on 2 February is an appropriate combination of initial settings to mimic actual spread in the Netherlands.

Performance of policies to prevent spread

There is ongoing research about the best policy to prevent spread. While some countries managed to effectively reduce spread, other countries were quickly overwhelmed. The Netherlands can be seen as one of the overwhelmed countries, being (at day of writing) ranked 75th by the country ranking of the Australian Lowy Institute. This institute compares performance of 98 countries with help of an own developed index. The relatively bad performance of the Netherlands in comparison to other countries could partially be due to population density, since countries with a smaller population generally have an advantage to deal with this kind of global crisis (Pandey, 2021). The Lowy Institute found that differences in population size between countries revealed the greatest difference in experiences with countries (Lowy Institute, n.d.). Unfortunately, this factor cannot be influenced with help of measures. Another cause of the low performance by the Netherlands could be timing, because timing of implementation can drastically influence the impact of measures (Haug et al., 2020). Or a lack of testing capacity, since it is important to combine measures that prevent spread of COVID-19 with a sufficient testing policy and contact tracing of infected cases to end the pandemic (Giordano et al., 2020). Below, our findings for strictness of measures, timing of implementation, and for testing capacity are compared to findings in literature.

- **Strictness of measures:** We observe that strictness of measures influences performance when comparing performance of mitigation and curbing type 1 to performance of curbing type 2 and curbing type 3. Mitigation and curbing type 1 determine their policy according to the signal values from the route map of the Dutch government. These policies both show higher values of spread than curbing type 2 and curbing type 3, and at the same time have a higher number of days with strict measures. However, mitigation and curbing type 1 apply more measures from the serious and very serious risk level, while curbing type 2 and curbing type 3 only apply lockdown measures. This indicates that signal values in the route map of the government are not sufficient to efficiently reduce spread, which supports the assumptions of Red Team. Red Team states that the signal values in the alarming, serious, and very serious risk level are not sufficient to come to the alert risk level quickly (RedTeam, 2020). Consequently, measures implemented in the risk levels alarming, serious, and very serious of the route map are too soft to effectively reduce spread.
- **Timing of measure implementation:** When comparing performance of policies, we see that implementation of strict measures is useful to reduce spread of COVID-19 effectively, especially when they are implemented early. This conforms to the findings of Haug et al. (2020), who suggest that one measure can have a drastically different impact when implemented earlier or later, and the findings proposed in an article from the NRC (a Dutch newspaper), who state among others that a quick and aggressive approach is effective (Berkhout, 2021). We observe timing to influence performance in our findings when we compare performance of curbing type 2 and curbing type 3. Curbing type 2 leads to both a lower number of days with strict measures and lower values for spread than curbing type 3. Curbing type 2 implements lockdown measures when the number of confirmed cases rises above 870 per day and soft measures when the number of confirmed cases declines below 870 per day, whereas curbing type 3 implements lockdown measures and soft measures in sequence every two weeks. This supports another statement of Red Team, who state that working with signal values from the number of confirmed cases works better to prevent spread than working with numbers of hospital admissions and IC occupations. Due to quicker availability of data for confirmed cases, this leads to better timing of measure implementation.
- **Testing capacity:** In accordance with our findings in literature, we observe that testing capacity has major impact on spread. One study found that testing is one of the most effective ways to manage the pandemic, yet it requires a large testing capacity and consequently many working hours. It is therefore important to combine a large testing capacity with other measures (Cui, Ni, & Shen, 2021). Several other studies support that it is important to combine measures that prevent spread of COVID-19 with a sufficient testing policy and with contact tracing of infected cases, to end the pandemic (Giordano et al., 2020). Contact tracing and extensive testing are particularly effective when combined with quarantine (Girum, Lentiro, Geremew, Migora, & Shewamare, 2020). This is in accordance with our findings the sensitivity analysis. While a high testing capacity shows to be very effective, we observe that a high testing capacity with a low fraction of infected cases entering quarantine is considerably less effective (Appendix O). There remains uncertainty about the magnitude of the effectiveness of testing still, but combining early detection, early implementation of quarantine and other measures are important to keep spread low (Nussbaumer-Streit et al., 2020)(Fang, Nie, & Penny, 2020).

6.3 Limitations

There remain some limitations in this study:

- Not all measures that are included in the route map of the government are included in our analysis (e.g. sports or the newly implement curfew). Some measures that we do not include might show their impact in the regression model with another measure that we do take into account, of which we believe facemasks to be the most apparent example.
- The number of tests on a day is not included in our prediction model for the effective reproduction rate, yet we did find a statistically significant relation between the two. Unfortunately, we are not able to substantiate the impact of the number of tests and therefore excluded the factor from our prediction model. In our current model, the measure facemasks shows a relatively high positive effect on the effective reproduction rate. We believe this factor might express some part of the impact previously expressed by the number of tests. We think a reason for this might be that the regression model wants to express an increase in contagiousness of the virus between February and December, or to express that people get tired of the pandemic, and uses the number of tests or the measure facemasks to express this in the model. Since both the number of tests per day and strictness of the measure facemasks (almost) only increase between February and December.
- In our regression model, we use assigned values 1, 2, 3, 4, or 5 to express the measures school openings, catering service openings, event allowance, and facemasks. We believe that this expression of government measures can be improved to obtain a better prediction model.
- We are aware that due to transformations of the virus, our findings might not be totally accurate anymore for future outbreaks of COVID-19 in the Netherlands. This is because new, more contagious variants of the virus are rising in the Netherlands and the same strict measures have less effect on spread. However, we believe that, at moment of writing, our findings for a remain valuable. Especially for the testing capacity and quarantine fraction.
- Lastly, clear communication is a vital part of the policy to prevent spread of COVID-19 (Anderson et al., 2020; Berkhout, 2021). The Dutch government has not always been clearly communicating with the Dutch population (e.g. when facemasks were recommended at some places but very few people actually used them at the beginning). Clear orders of the government seem to be most effective to reduce the effective reproduction rate for approximately two weeks, while recommendations seem to be not effective at all. Also Red Team believes clear communication to have an important impact on the effectiveness of the policy to reduce spread. We did not study the effect of clear communication, but we believe that clear communication can affect performance of policies considerably.

6.4 Recommendations

While the virus of COVID-19 is changing with upcoming new variants, we believe the following two recommendations remain important to make impact of measures on spread in the Netherlands most effective. Combining recommendation 1 and 2 is key to limit damage to public health and economy.

1) Implement strict measures early

Strict measures are most effective when implemented early to prevent a quick rise in the number of infected cases. We recommend to use signal values of the number of confirmed cases per day to base the policy on, and to lower the signal values per risk level in the route map of the government.

2) Invest in a high testing capacity and high quarantine fraction simultaneously

Combining a high number of tests with a high quarantine fraction shows to be very effective to prevent spread. A high number of tests with a low quarantine fraction is considerably less effective, which can even affect effectiveness of the implemented policy.

References

Aleva, F. E., van Mourik, L., Broeders, M. E. A. C., Paling, A. J., & de Jager, C. P. C. (2020). COVID-19 in critically ill patients in North Brabant, the Netherlands: Patient characteristics and outcomes. *Journal of Critical Care*, *60*, 111–115. https://doi.org/10.1016/j.jcrc.2020.08.001

Anderson, R. M., Heesterbeek, H., Klinkenberg, D., & Hollingsworth, T. D. (2020). How will country-based mitigation measures influence the course of the COVID-19 epidemic? *The Lancet*, *395*(10228), 931–934. https://doi.org/10.1016/S0140-6736(20)30567-5

Berkhout, K. (2021, January 31). Nut van maatregelen: 'Europa wachtte te lang.' *NRC*. Retrieved from https://www.nrc.nl/nieuws/2021/01/31/handel-bij-een-pandemie-snel-en-krachtig-a4029959

Block, P., Hoffman, M., Raabe, I. J., Dowd, J. B., Rahal, C., Kashyap, R., & Mills, M. C. (2020). Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nature Human Behaviour*, *4*(6), 588–596. https://doi.org/10.1038/s41562-020-0898-6

Byrne, A. W., McEvoy, D., Collins, A. B., Hunt, K., Casey, M., Barber, A., … More, S. J. (2020). Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases. *BMJ Open*, *10*(8), e039856. https://doi.org/10.1136/bmjopen-2020-039856

Cartaud, A., Quesque, F., & Coello, Y. (2020). Wearing a face mask against Covid-19 results in a reduction of social distancing. *PLOS ONE*, *15*(12), e0243023. https://doi.org/10.1371/journal.pone.0243023

CBS. (2020, October 1). *10 duizend coronadoden tijdens eerste golf van de pandemie*. Retrieved from https://www.cbs.nl/nl-nl/nieuws/2020/40/10-duizend-coronadoden-tijdens-eerste-golf-van-de-pandemie

Cui, Y., Ni, S., & Shen, S. (2021). A network-based model to explore the role of testing in the epidemiological control of the COVID-19 pandemic. *BMC Infectious Diseases*, *21*(1), 58. https://doi.org/10.1186/s12879-020-05750-9

Datagraver. (2020, October 3). Verloop COVID-19 Nederland. Retrieved October 14, 2020, from https://www.datagraver.com/case/verloop-covid-19-nederland

de Beus, T., Rooze, M., & van Bree, T. (2020, August 14). *Drie belangrijke ontwikkelingen van de tweede coronagolf om in de gaten te houden | Pointer*. Retrieved from https://pointer.kro-ncrv.nl/artikelen/drie-belangrijke-ontwikkelingen-van-de-tweede-coronagolf-om-in-de-gaten-te-houden

Dupuy, L. (2020, September 11). Teststraat Schiphol tijdelijk gesloten door gebrek aan testcapaciteit. *NRC*. Retrieved from https://www.nrc.nl/nieuws/2020/09/11/coronablog-11-september-a4011604#upd20200911210059

ECDC. (2020). *Coronavirus disease 2019 (COVID-19) in the EU/EEA and the UK-eleventh update: resurgence of cases*.

Fang, Y., Nie, Y., & Penny, M. (2020). Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis. *Journal of Medical Virology*, *92*(6), 645–659. https://doi.org/10.1002/jmv.25750

Feenstra, W., & Effting, M. (2020, August 28). Testcapaciteit corona moet omhoog, maar laboratoria kampen nu al met tekorten. *De Volkskrant*. Retrieved from https://www.volkskrant.nl/nieuws-achtergrond/testcapaciteit-corona-moet-omhoog-maar-laboratoria-kampen-nu-al-met-tekorten~bcf99252/

Fisman, D., Greer, A., Hillmer, M., O'Brien, S., Drews, S. J., & Tuite, A. (2020). *COVID-19 Case-Age Distribution: Correction for Differential Testing by Age*. Retrieved from https://medrxiv.org/cgi/content/short/2020.09.15.20193862

Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., & Colaneri, M. (2020). Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*, *26*(6), 855–860. https://doi.org/10.1038/s41591-020-0883-7

Girum, T., Lentiro, K., Geremew, M., Migora, B., & Shewamare, S. (2020, December 1). Global strategies and effectiveness for COVID-19 prevention through contact tracing, screening, quarantine, and isolation: a systematic review. *Tropical Medicine and Health*, Vol. 48, p. 91. https://doi.org/10.1186/s41182-020-00285-w

Google.nl. (2020). COVID-19 Community Mobility Reports. Retrieved November 6, 2020, from https://www.google.com/covid19/mobility/data_documentation.html?hl=nl

Haug, N., Geyrhofer, L., Londei, A., Dervic, E., Desvars-Larrive, A., Loreto, V., … Klimek, P. (2020). Ranking the effectiveness of worldwide COVID-19 government interventions. *Nature Human Behaviour*, *4*(12), 1303–1312. https://doi.org/10.1038/s41562-020-01009-0

He, W., Yi, G. Y., & Zhu, Y. (2020). Estimation of the basic reproduction number, average incubation time, asymptomatic infection rate, and case fatality rate for COVID-19: Meta-analysis and sensitivity analysis. *Journal of Medical Virology*, jmv.26041. https://doi.org/10.1002/jmv.26041

Heerkens, H., & van Winden, A. (2017). Solving Managerial Problems Systematically. Retrieved August 20, 2020, from Noordhoff Uitgevers website: https://web.b.ebscohost.com/ehost/ebookviewer/ebook/bmxlYmtfXzE0OTE0NDlfX0FO0?sid=395b6ef6-220b-4985-9055-3f7d8499132e@sessionmgr101&vid=0&format=EB&rid=1

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *Springer Texts in Statistics An Introduction to Statistical Learning*. Retrieved from http://www.springer.com/series/417

Katri Manninen. (2020, October 8). Typical Progress of Covid-19. Retrieved October 22, 2020, from https://katrimanninen.net/covid19progress/

Knight, J., & Mishra, S. (2020). Estimating effective reproduction number using generation time versus serial interval, with application to COVID-19 in the Greater Toronto Area, Canada. *Infectious Disease Modelling*, *5*, 889–896. https://doi.org/10.1016/j.idm.2020.10.009

KNMI. (n.d.). Vochtigheid. Retrieved September 17, 2020, from https://www.knmi.nl/kennis-en-datacentrum/uitleg/vochtigheid

KNMI. (2020). Daggegevens van het weer in Nederland. Retrieved September 17, 2020, from http://projects.knmi.nl/klimatologie/daggegevens/selectie.cgi

Linka, K., Peirlinck, M., & Kuhl, E. (2020). The reproduction number of COVID-19 and its correlation with public health interventions. *Computational Mechanics*, *66*(4), 1035–1050. https://doi.org/10.1007/s00466-020-01880-8

Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E., & Getz, W. M. (2005). *Superspreading and the effect of individual variation on disease emergence*. https://doi.org/10.1038/nature04153

Lowy Institute. (n.d.). Covid Performance. Retrieved February 4, 2021, from https://interactives.lowyinstitute.org/features/covid-performance/#rankings

Marshall, D. A., Burgos-Liz, L., Eng, I., Ijzerman, M. J., Crown, W., Padula, W. V, … Osgood, N. D. (n.d.). *Selecting a Dynamic Simulation Modeling Method for Health Care Delivery Research-Part 2: Report of the ISPOR Dynamic Simulation Modeling Emerging Good Practices Task Force Background to the Task Force*. https://doi.org/10.1016/j.jval.2015.01.006

Mecenas, P., da Rosa Moreira Bastos, R. T., Rosário Vallinoto, A. C., & Normando, D. (2020). Effects of temperature and humidity on the spread of COVID-19: A systematic review. *PLoS ONE*, *15*(9 September). https://doi.org/10.1371/journal.pone.0238339

Metatopos. (2020). Database for topographical data. Retrieved September 18, 2020, from https://www.metatopos.eu/provincies_eu.php

NICE. (2020). Covid-19 op de IC. Retrieved October 23, 2020, from https://www.stichting-nice.nl/covid-19-op-de-ic.jsp

Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., … Agha, R. (2020, June 1). The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *International Journal of Surgery*, Vol. 78, pp. 185–193. https://doi.org/10.1016/j.ijsu.2020.04.018

Nikolai, L. A., Meyer, C. G., Kremsner, P. G., & Velavan, T. P. (2020). Asymptomatic SARS Coronavirus 2 infection: Invisible yet invincible. *International Journal of Infectious Diseases : IJID : Official Publication of the International Society for Infectious Diseases*. https://doi.org/10.1016/j.ijid.2020.08.076

NOS. (2020, September 22). *GGD: tienduizenden mensen kunnen niet getest worden op corona*. Retrieved from https://nos.nl/artikel/2349455-ggd-tienduizenden-mensen-kunnen-niet-getest-worden-op-corona.html

NU.nl. (2020a, July 28). *Kabinet positief over bestrijding corona, testcapaciteit is wel een probleem*. Retrieved from https://www.nu.nl/binnenland/6067261/kabinet-positief-over-bestrijding-corona-testcapaciteit-is-wel-een-probleem.html

NU.nl. (2020b, July 28). *We laten ons meer testen, maar kan de GGD dat nog wel aan?* Retrieved from https://www.nu.nl/dit-wordt-het-nieuws/6067235/we-laten-ons-meer-testen-maar-kan-de-ggd-dat-nog-wel-aan.html

NU.nl. (2020c, September 22). *Van Dissel over toename besmettingen: “We staan op een kantelpunt.”* Retrieved from https://www.nu.nl/coronavirus/6079108/van-dissel-over-toename-besmettingen-we-staan-op-een-kantelpunt.html

NU.nl. (2020d, September 23). *Toon in coronadebat slaat om: “Kabinetsbeleid staat onder curatele.”* Retrieved from https://www.nu.nl/coronavirus/6079222/toon-in-coronadebat-slaat-om-kabinetsbeleid-staat-onder-curatele.html

NU.nl. (2020e, September 23). *We zitten in de tweede golf, wat betekent dat voor de economie?* Retrieved from https://www.nu.nl/economie/6079132/we-zitten-in-de-tweede-golf-wat-betekent-dat-voor-de-economie.html

NU.nl. (2020f, September 29). *Zo komt het sterftecijfer onder coronapatiënten tot stand*. Retrieved from https://www.nu.nl/uitleg-over-het-coronavirus/6080383/zo-komt-het-sterftecijfer-onder-coronapatienten-tot-stand.html

NU.nl. (2020g, December 1). *Dit weten we over de werking van mondkapjes en (schijn)veiligheid*. Retrieved from https://www.nu.nl/uitleg-over-het-coronavirus/6093836/dit-weten-we-over-de-werking-van-mondkapjes-en-schijnveiligheid.html

Nussbaumer-Streit, B., Mayr, V., Dobrescu, A. I., Chapman, A., Persad, E., Klerings, I., … Gartlehner, G. (2020). Quarantine alone or in combination with other public health measures to control COVID-19: a rapid review. *Cochrane Database of Systematic Reviews*, *2020*(9). https://doi.org/10.1002/14651858.CD013574.pub2

Pandey, S. (2021, January 28). New Zealand, Vietnam top COVID performance ranking; U.S., UK languish. *Reuters*. Retrieved from https://www.reuters.com/article/us-health-coronavirus-australia-lowy/new-zealand-vietnam-top-covid-performance-ranking-u-s-uk-languish-idUSKBN29W311

Pascarella, G., Strumia, A., Piliego, C., Bruno, F., Del Buono, R., Costa, F., … Agrò, F. E. (2020, August 1). COVID-19 diagnosis and management: a comprehensive review. *Journal of Internal Medicine*, Vol. 288, pp. 192–206. https://doi.org/10.1111/joim.13091

Peirlinck, M., Linka, K., Sahli Costabal, F., Bhattacharya, J., Bendavid, E., Ioannidis, J. P. A., & Kuhl, E. (2020). *Visualizing the invisible: The effect of asymptomatic transmission on the outbreak dynamics of COVID-19*. https://doi.org/10.1101/2020.05.23.20111419

Qiu, J., Shen, B., Zhao, M., Wang, Z., Xie, B., & Xu, Y. (2020). A nationwide survey of psychological distress among Chinese people in the COVID-19 epidemic: implications and policy recommendations. *General Psychiatry*, *33*(2), e100213. https://doi.org/10.1136/gpsych-2020-100213

Ranzijn, A. (2020, March 24). Cijfers RIVM: welke leeftijdscategorie is het zwaarst getroffen? *Het Parool*. Retrieved from https://www.parool.nl/nieuws/cijfers-rivm-welke-leeftijdscategorie-is-het-zwaarst-getroffen~b06453055/

RedTeam. (2020, October 22). Adviezen COVID-19 NL. Retrieved January 11, 2021, from https://www.c19redteam.nl/adviezen/

Rijksoverheid. (2020a). Dashboard Coronavirus. Retrieved November 9, 2020, from https://coronadashboard.rijksoverheid.nl/landelijk/positief-geteste-mensen

Rijksoverheid. (2020b). Documenten Coronavirus COVID-19. Retrieved February 19, 2021, from https://www.rijksoverheid.nl/onderwerpen/coronavirus-covid-19/documenten?trefwoord=&startdatum=&einddatum=&onderwerp=Alle+onderwerpen&onderdeel=Alle+ministeries&type=Alle+documenten

Rijksoverheid. (2020c). *Letterlijke tekst persconferentie 09-03-2020*. Retrieved from https://www.rijksoverheid.nl/onderwerpen/coronavirus-covid-19/documenten/mediateksten/2020/03/09/letterlijke-tekst-persconferentie-minister-president-mark-rutte-en-directeur-jaap-van-dissel-rivm-na-afloop-van-crisisberaad-kabinet

Rijksoverheid. (2020d, January 31). 98e OMT advies deel 1 | Rapport. Retrieved February 2, 2021, from https://www.rijksoverheid.nl/documenten/rapporten/2021/01/31/98e-omt-advies-deel-1

Rijksoverheid. (2020e, May 27). *Newsupdate testing after 1 June*. Retrieved from https://www.rijksoverheid.nl/actueel/nieuws/2020/05/27/vanaf-1-juni-testen-mogelijk-voor-iedereen-met-milde-klachten

Rijksoverheid. (2020f, June 3). *Letterlijke tekst persconferentie 03-06-2020*. Retrieved from https://www.rijksoverheid.nl/onderwerpen/coronavirus-covid-19/documenten/mediateksten/2020/06/03/letterlijke-tekst-persconferentie-minister-president-rutte-en-minister-de-jonge-na-afloop-van-crisisberaad-kabinet

Rijksoverheid. (2020g, November 4). Routekaart coronamaatregelen. Retrieved January 11, 2021, from https://www.rijksoverheid.nl/documenten/publicaties/2020/10/13/risiconiveaus-en-maatregelen-covid-19

RIVM. (2020a). Actuele informatie over het coronavirus. Retrieved January 4, 2021, from https://www.rivm.nl/coronavirus-covid-19/actueel

RIVM. (2020b). Data Catalogus. Retrieved January 5, 2021, from https://data.rivm.nl/geonetwork/srv/dut/catalog.search#/search?isChild='false'&resultType=details&fast=index&_content_type=json&from=1&to=20&sortBy=relevance&any_OR__title=covid

RIVM. (2020c). Hoe rekenmodellen bijdragen aan de bestrijding van COVID-19. Retrieved January 4, 2021, from https://www.rivm.nl/coronavirus-covid-19/rekenmodellen

RIVM. (2020d). LCI richtlijn COVID-19. Retrieved August 20, 2020, from https://lci.rivm.nl/richtlijnen/covid-19

RIVM. (2020e). Modelling the spread of the coronavirus SARS-CoV-2. Retrieved January 25, 2021, from https://www.rivm.nl/en/novel-coronavirus-covid-19/modelling

RIVM. (2020f). Ontwikkeling COVID-19 in grafieken. Retrieved August 28, 2020, from https://www.rivm.nl/coronavirus-covid-19/grafieken

RIVM. (2020g). *Patiënt met COVID-19 in Delft*. Retrieved from https://www.rivm.nl/nieuws/patient-met-covid-19-in-delft

RIVM. (2020h). *Patient with novel coronavirus COVID-19 in the Netherlands*. Retrieved from https://www.rivm.nl/node/152811

RIVM. (2020i). *Tweede patiënt in Nederland met COVID-19*. Retrieved from https://www.rivm.nl/nieuws/tweede-patient-in-nederland-met-covid-19

RIVM. (2020j). Wekelijkse update epidemiologische situatie COVID-19 in Nederland. Retrieved January 4, 2021, from https://www.rivm.nl/coronavirus-covid-19/actueel/wekelijkse-update-epidemiologische-situatie-covid-19-in-nederland

RIVM. (2020k, April 10). *Regionale verschillen in de epidemie*. Retrieved from https://www.rivm.nl/nieuws/regionale-verschillen-in-epidemie

RIVM. (2020l, August 13). Testen op COVID-19. Retrieved September 21, 2020, from https://www.rivm.nl/coronavirus-covid-19/testen

RIVM. (2020m, December 28). *Verspreiding Britse coronavirusvariant in Nederland*. Retrieved from https://www.rivm.nl/nieuws/verspreiding-britse-coronavirusvariant-in-nederland

Rozhnova, G., Van Dorp, C. H., Bruijning-Verhagen, P., Bootsma, M. C. J., Van De Wijgert, J. H. H. M., Bonten, M. J. M., … Kretzschmar, E. (2020). Model-based evaluation of school- and non-school-related measures to control the COVID-19 pandemic. *MedRxiv*, 2020.12.07.20245506. https://doi.org/10.1101/2020.12.07.20245506

RTLnieuws. (2020a, June 4). *Eerste Nederlandse coronapatiënten hadden al klachten in januari*. Retrieved from https://www.rtlnieuws.nl/nieuws/artikel/5140371/corona-cijfers-besmettingen-uitbraak-nederland-eerste-officiele

RTLnieuws. (2020b, October 11). *Dit is het verschil tussen een intelligente en volledige lockdown*. Retrieved from https://www.rtlnieuws.nl/nieuws/nederland/artikel/5189608/lockdown-coronamaatregelen-coronabesmettingen-coronavirus-regels

Sakurai, A., Sasaki, T., Kato, S., Hayashi, M., Tsuzuki, S., Ishihara, T., … Doi, Y. (2020). Natural History of Asymptomatic SARS-CoV-2 Infection. *New England Journal of Medicine*, *383*(9), 885–886. https://doi.org/10.1056/nejmc2013020

Sciensano. (2020). Coronavirus Covid-19. Retrieved February 25, 2021, from https://covid-19.sciensano.be/nl

Statline. (2020, May 27). Bevolking op 1 januari en gemiddeld; geslacht, leeftijd en regio. Retrieved September 17, 2020, from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/03759ned/table?ts=1600347087739

Steyerberg, E. W. (2009). *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating (Statistics for Biology and Health) - EPDF.PUB*. Retrieved from https://epdf.pub/queue/clinical-prediction-models-a-practical-approach-to-development-validation-and-up83807.html

Tabatabaeizadeh, S. A. (2021, December 1). Airborne transmission of COVID-19 and the role of face mask to prevent it: a systematic review and meta-analysis. *European Journal of Medical Research*, Vol. 26, p. 1. https://doi.org/10.1186/s40001-020-00475-6

Translink. (2020). Impact van het coronavirus op het gebruik van openbaar vervoer in Nederland. Retrieved November 6, 2020, from https://www.translink.nl/library

van de Klundert, M. (2020, June 5). Eerste testweek: half miljoen telefoontjes, en toch maar 40.000 tests. *NOS*. Retrieved from https://nos.nl/artikel/2336315-eerste-testweek-half-miljoen-telefoontjes-en-toch-maar-40-000-tests.html

van Kempen, J., & Soetenhorst, B. (2020, August 1). Het aantal besmettingen groeit fors. Wanneer is het tijd om in te grijpen? *Het Parool*. Retrieved from https://www.parool.nl/nederland/het-aantal-besmettingen-groeit-fors-wanneer-is-het-tijd-om-in-te-grijpen~b69825c2/

Vensim. (n.d.). Online courses. Retrieved February 8, 2021, from https://vensim.com/vensim-online-courses/

Venticare. (2020). NVIC geeft dagelijks inzicht in IC-capaciteit voor COVID-19. Retrieved October 23, 2020, from https://venticare.nl/nieuws/nvic-geeft-dagelijks-inzicht-in-ic-capaciteit-voor-covid19.html

Viner, R. M., Russell, S. J., Croker, H., Packer, J., Ward, J., Stansfield, C., … Booy, R. (2020, May 1). School closure and management practices during coronavirus outbreaks including COVID-19: a rapid systematic review. *The Lancet Child and Adolescent Health*, Vol. 4, pp. 397–404. https://doi.org/10.1016/S2352-4642(20)30095-X

Wallinga, J., & Teunis, P. (2004). Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *American Journal of Epidemiology*, *160*(6), 509–516. https://doi.org/10.1093/aje/kwh255

Warmedagen. (2020). warmte en het coronavirus. Retrieved September 17, 2020, from https://www.warmedagen.be/warmte-en-het-coronavirus

WHO. (2020a). *Estimating mortality from COVID-19*. Retrieved from https://www.who.int/news-room/commentaries/detail/estimating-mortality-from-covid-19

WHO. (2020b). Timeline COVID-19. Retrieved August 18, 2020, from https://www.who.int/emergencies/diseases/novel-coronavirus-2019/interactive-timeline#event-66

WHO. (2020c). *Transmission of SARS-CoV-2*. Retrieved from https://www.who.int/news-room/commentaries/detail/transmission-of-sars-cov-2-implications-for-infection-prevention-precautions

WHO. (2020d, July 9). *Coronavirus disease (COVID-19): How is it transmitted?* Retrieved from https://www.who.int/news-room/q-a-detail/coronavirus-disease-covid-19-how-is-it-transmitted

WHO. (2020e, September 17). Dashboard. Retrieved September 18, 2020, from https://covid19.who.int/table

Wigginton, K. R., & Boehm, A. B. (2020). Environmental Engineers and Scientists Have Important Roles to Play in Stemming Outbreaks and Pandemics Caused by Enveloped Viruses. *Cite This: Environ. Sci. Technol*, *54*, 3739. https://doi.org/10.1021/acs.est.0c01476

Worldometers. (2020a, May 13). Coronavirus Age, Sex, Demographics (COVID-19). Retrieved September 25, 2020, from https://www.worldometers.info/coronavirus/coronavirus-age-sex-demographics/

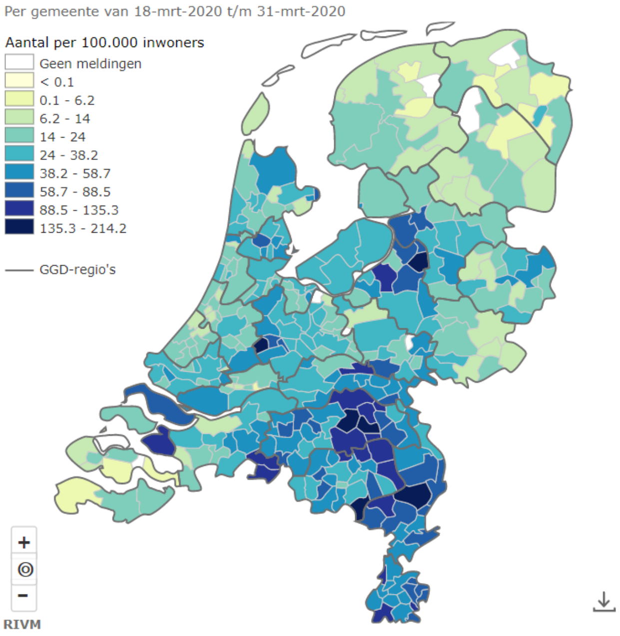
Worldometers. (2020b, May 14). Mortality Rate COVID-19. Retrieved September 21, 2020, from https://www.worldometers.info/coronavirus/coronavirus-death-rate/#days

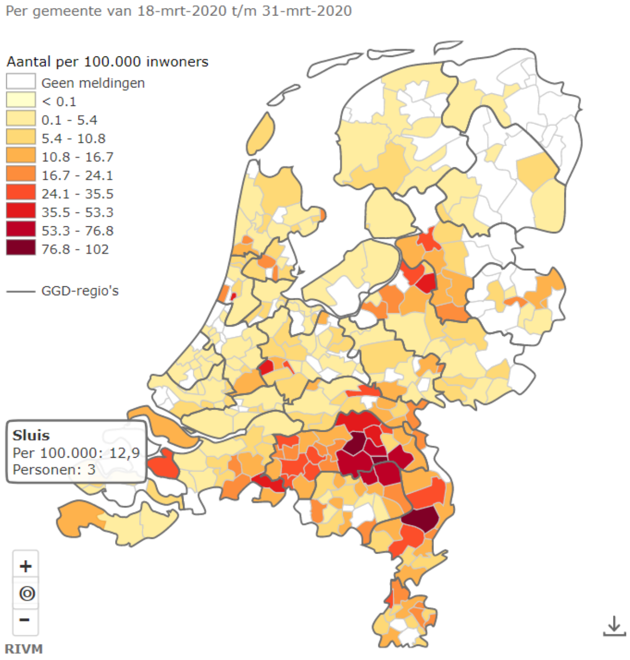
Wu, Z., & McGoogan, J. M. (2020, April 7). Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases from the Chinese Center for Disease Control and Prevention. *JAMA - Journal of the American Medical Association*, Vol. 323, pp. 1239–1242. https://doi.org/10.1001/jama.2020.2648

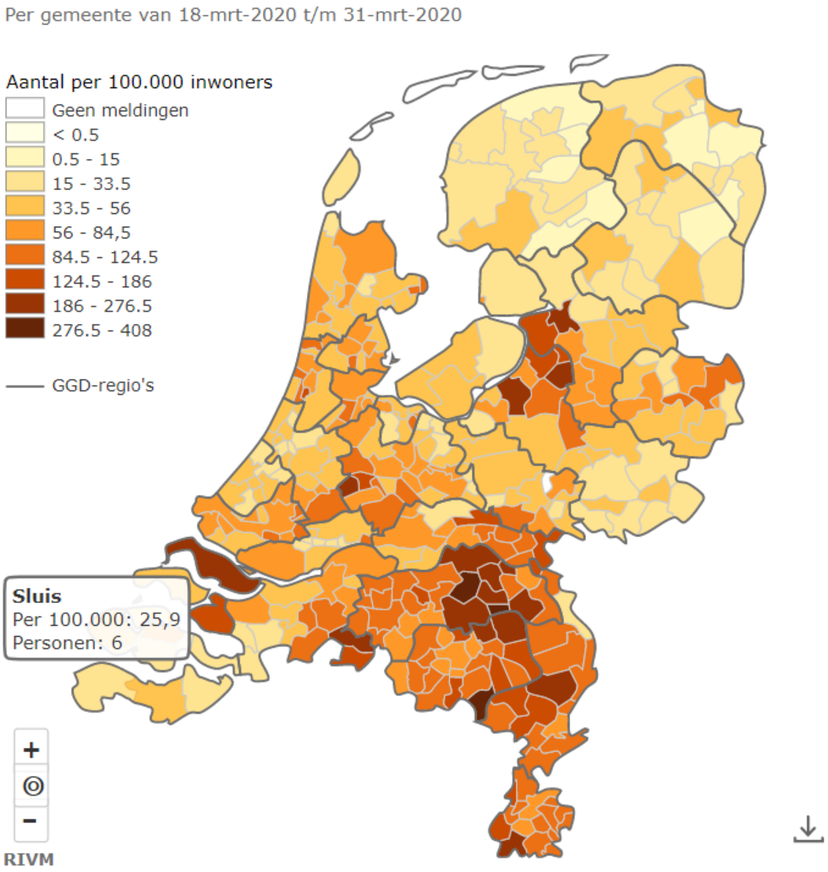
Appendix

A. Spread of COVID-19 provinces

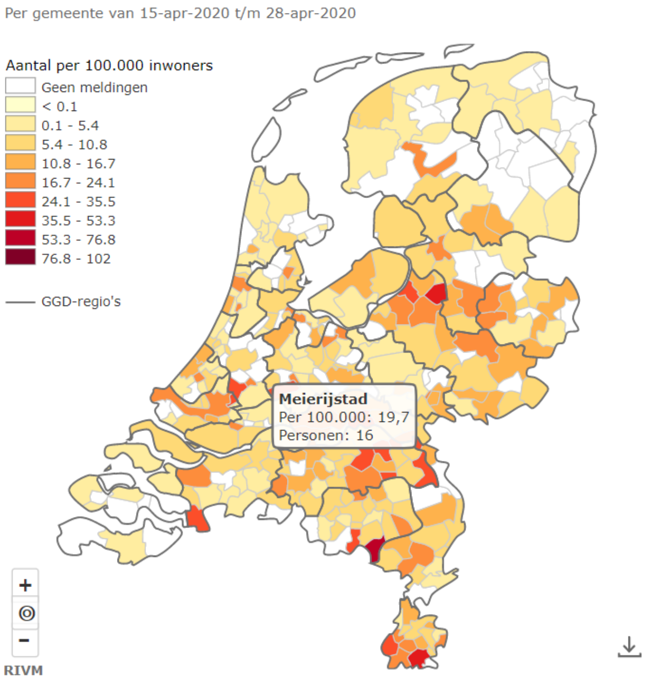
Start first peak

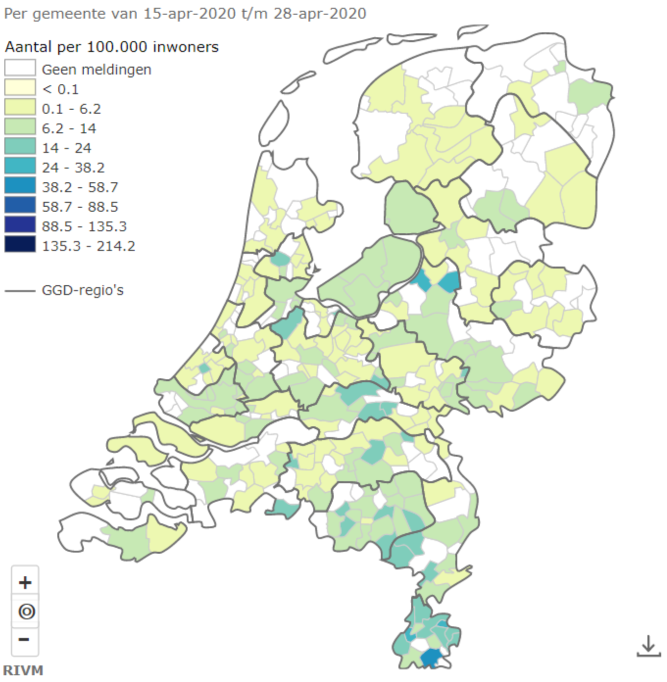


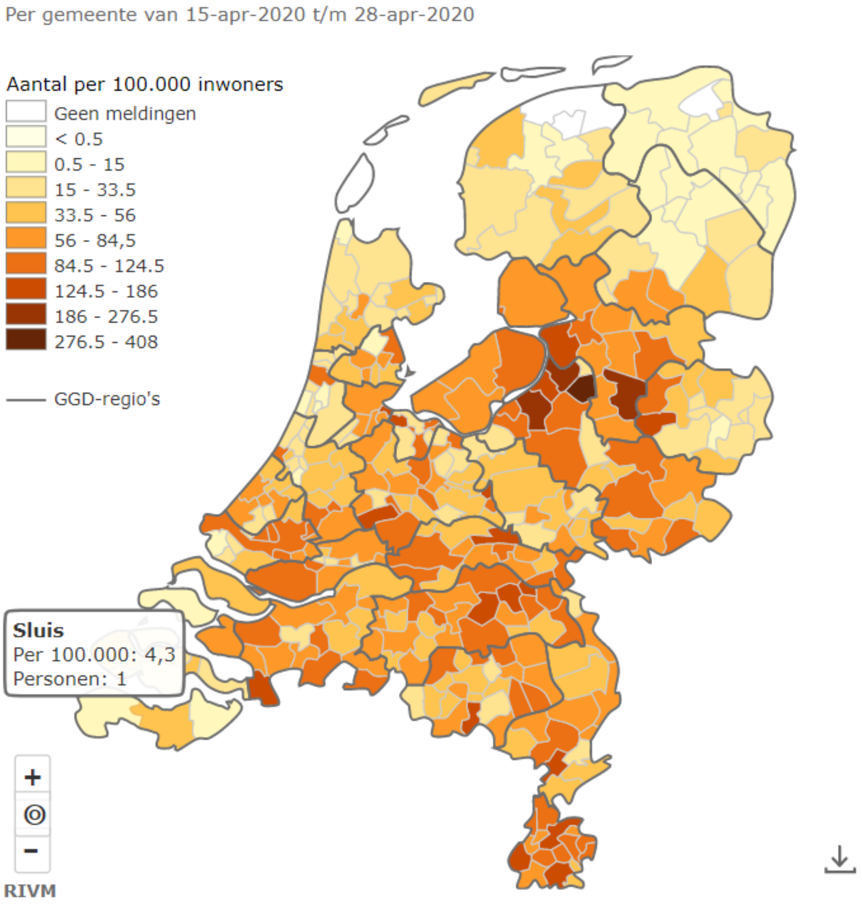




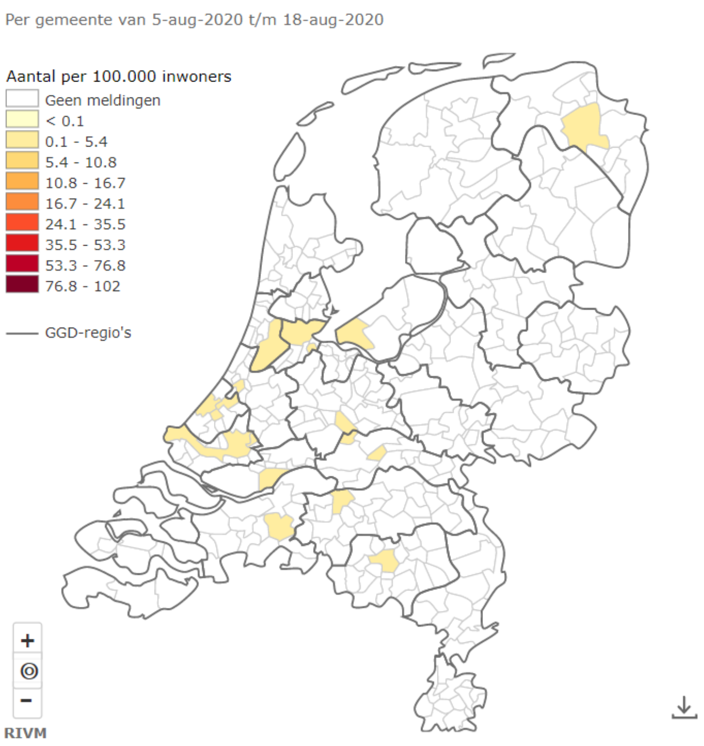
End first peak

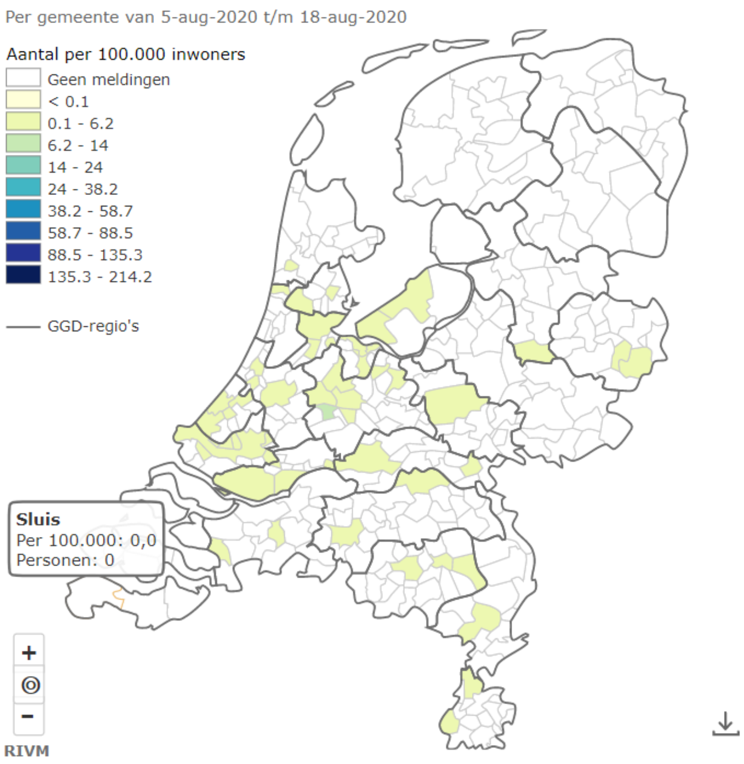


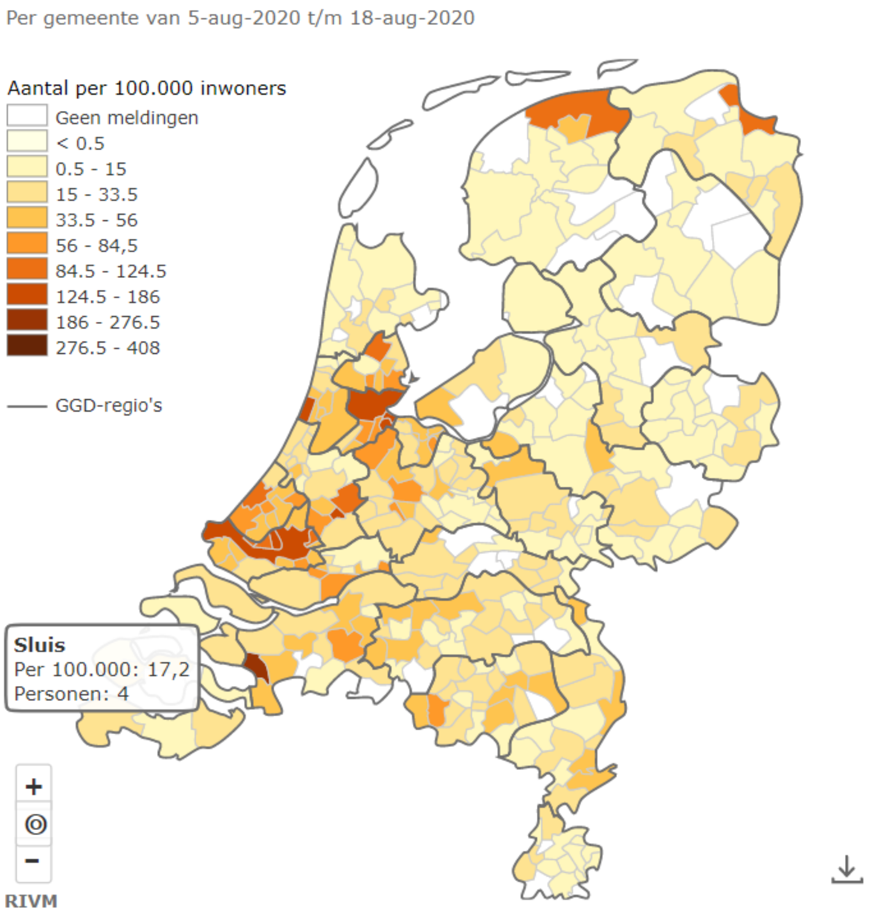




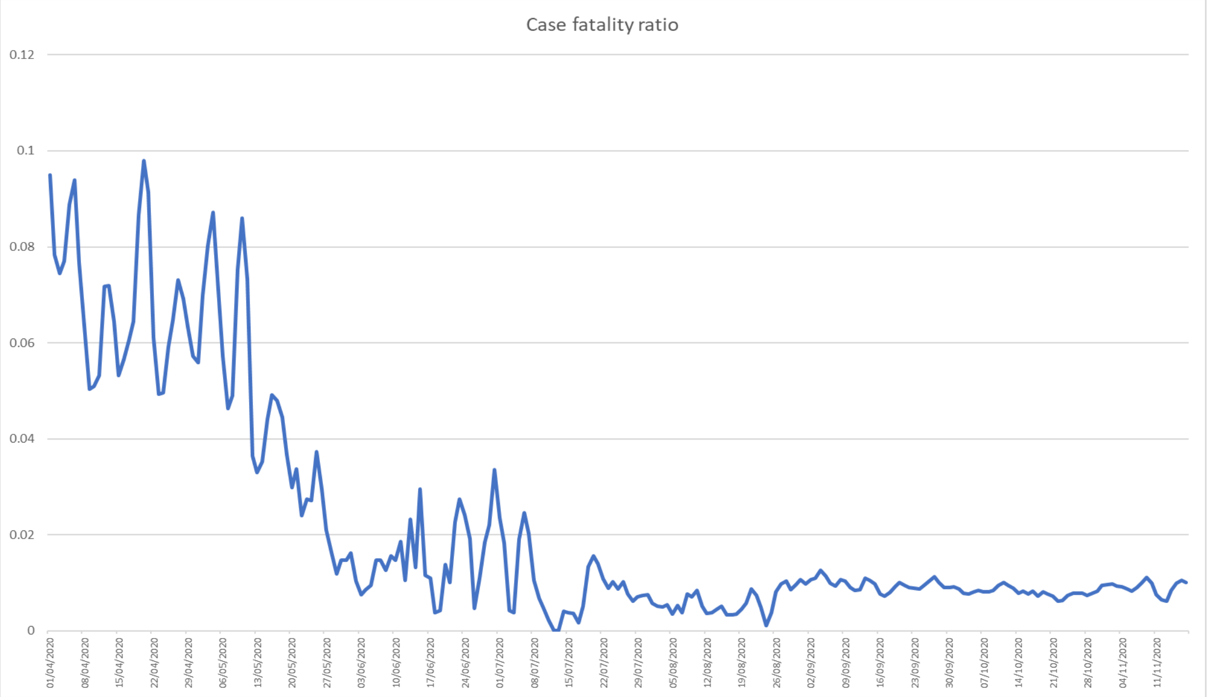
Start second peak







B. Development of the case fatality rate in the Netherlands



C. Population density versus spread per province

Province	Share of infections per province	Share of infections per province (adjusted)	Share of hospitalizations per province	Share of hospitalizations per province (adjusted)	Share of per deaths province	Share of deaths per province (adjusted)
Groningen	0.01	0.03	0.01	0.01	0.00	0.01
Friesland	0.01	0.04	0.01	0.03	0.01	0.03
Drenthe	0.01	0.03	0.01	0.03	0.01	0.02
Overijssel	0.04	0.08	0.05	0.07	0.05	0.08
Flevoland	0.02	0.04	0.02	0.04	0.02	0.03
Gelderland	0.10	0.16	0.12	0.18	0.11	0.16
Utrecht	0.08	0.06	0.08	0.05	0.07	0.05
Noord-Holland	0.20	0.12	0.14	0.08	0.13	0.07
Zuid-Holland	0.30	0.14	0.20	0.08	0.22	0.09
Zeeland	0.01	0.04	0.01	0.03	0.01	0.03
Noord-Brabant	0.15	0.19	0.22	0.25	0.25	0.28
Limburg	0.06	0.08	0.13	0.14	0.12	0.14

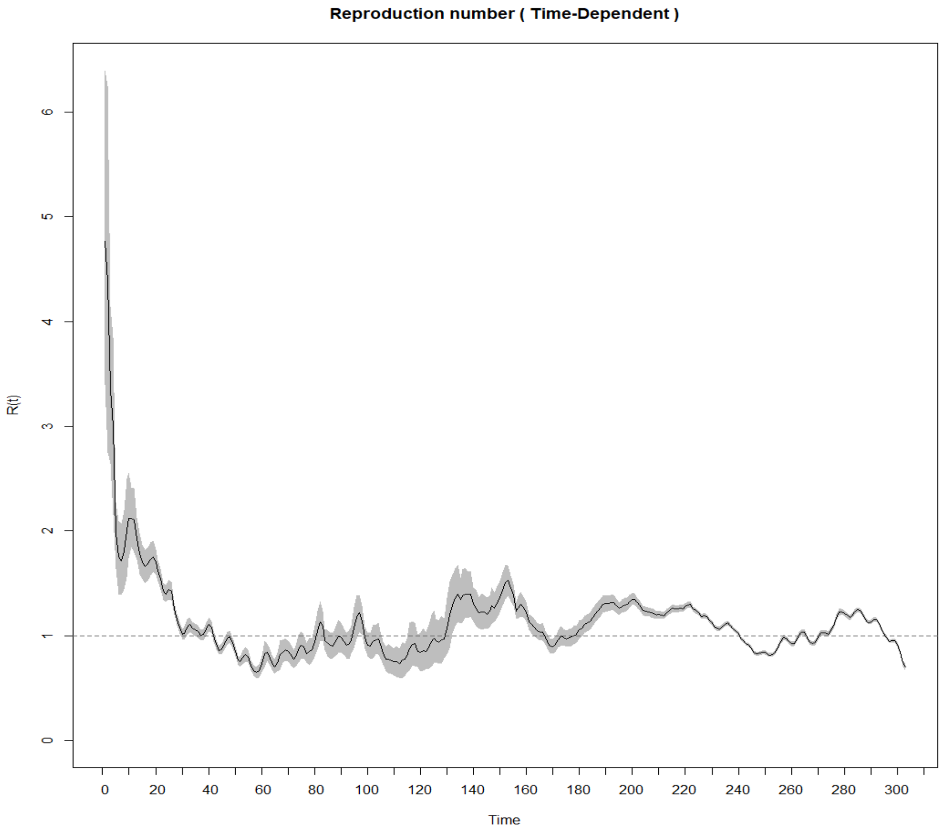
D. Defining values of measures

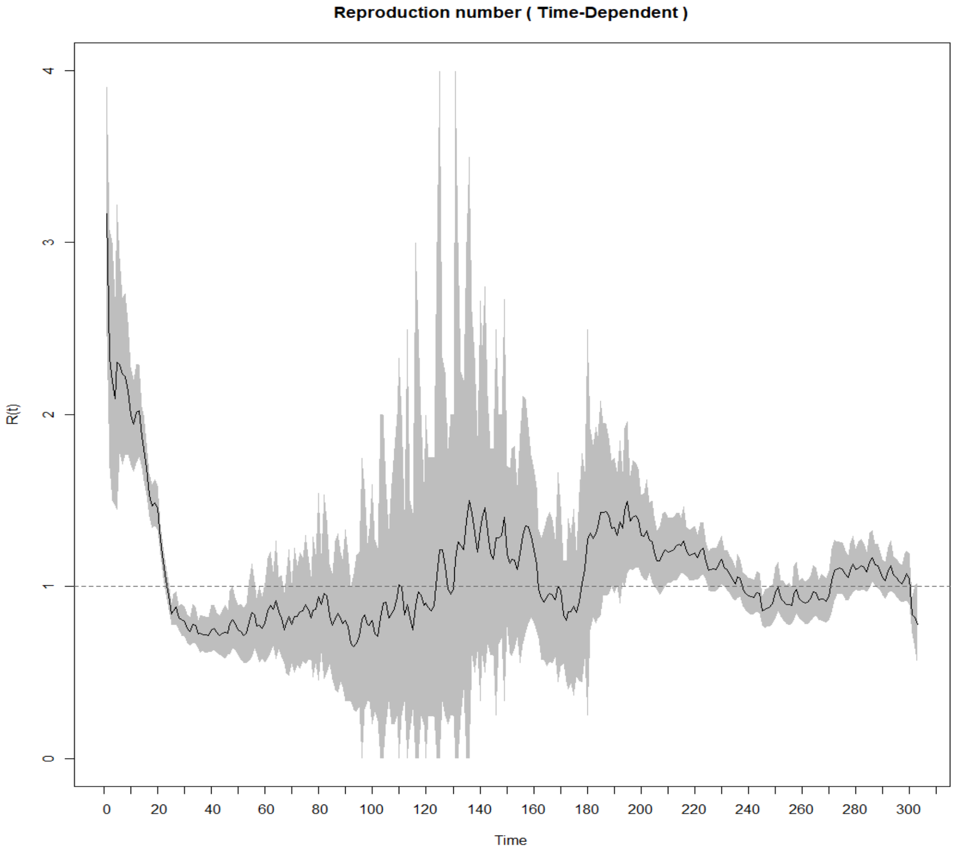
Measure	Value	Reasoning	Specification
School openings	1	No measure	Normal education. Fully physical.
	3	From 12 to 15 March, universities suspend physical education.	Partially physical, partially non-physical education / Approx. half of schools opened.
	5	After 15 March all schools close until 12 May, assigned value is 5 (lockdown measure).	All schools closed / only online education.
	4	At 12 May primary schools open.	Mostly non-physical education / most schools closed.
	3	At 2 June secondary schools also open.	Partially physical, partially non-physical education / Approx. half of schools opened.
	2	At 15 June universities are allowed to partially open for practical education and exams.	Mostly physical education / Most schools open.
	5	At 8 July all schools close due to summer holidays, the value becomes 5 (hard measure) since no contact anymore at schools.	All schools closed / only online education.
	3	At 17 August some primary and secondary schools open again.	Partially physical, partially non-physical education / Approx. half of schools opened.
	2	On 1 September universities are also open again and value 2 is assigned until 16 December.	Mostly physical education / Most schools open.
	4	Two days before the start of the Christmas holidays, the primary and secondary schools closed. This was on 16 December	Mostly non-physical education / most schools closed.
	5	On 18 December all schools and universities closed due to the Christmas holidays and the stricter government measures.	All schools closed / only online education.
Catering services openings	1	No measure; until 16 March	All catering services normally opened.
	5	Every catering service is closed until 1 June	All catering services closed.
	3	On 1 June catering services opened with moderate capacity/strict rules.	All catering services opened, closing at 12 PM.
	4	On 28 September catering services have to close at 22 o'clock	All catering services opened, closing at 10 PM.
	5	On 13 October all catering services close	All catering services closed.
Event allowance	1	No measure	All events allowed.
	3	At 12 March events and gatherings with more than 100 persons are cancelled until 23 March.	Big events prohibited.
	5	After 23 March all events are prohibited. Until 1 June.	All events and gatherings are prohibited.
	4	From 1 June until 1 July, gatherings with a maximum of 30 people are allowed and outside activities may find place with a maximum of 250 people.	Events prohibited and gatherings with maximum number of people.
	3	From 1 July inside activities do not have a maximum allowed number of people anymore. Yet people do have to stay on distance from each other.	Big events prohibited.

	4	On 20 September measures are sharper. People were only allowed to meet with maximum 50 people (and from 1 October a maximum of 30 persons inside and 40 persons outside).	Events prohibited and gatherings with maximum number of people.
	5	On 13 October all events prohibited.	All events and gatherings are prohibited. No maximum.
Facemasks	1	No measure; until 1 June	No facemasks.
	2	From 1 June wearing facemasks was obliged in public transport.	Facemasks in public transport only.
	3	From 28 September in some regions people were urged to wear facemasks	Urge to wear facemasks.
	4	From 13 October the government strongly urged people to wear facemasks in all public spaces.	Strong urge to wear facemasks in all public spaces.
	5	On 1 December, wearing facemasks became mandatory.	Facemasks mandatory in all public spaces.

E. Calculated effective reproduction rate in R

Above figure is based on the number of infections, beneath based on the number of hospitalizations.





F. Input data for multiple regression

Parameter	**Values**	**Value range**
Reproduction rate	*Numeric*	*Min: 0.600* *Max: 3.270* *Mean: 1.085*
Relative humidity	*Percentage*	*Min: 38* *Max: 96* *Mean: 75*
Average temperature	*Degrees Celsius*	*Min: 0.9* *Max: 27.0* *Mean: 12.9*
Maximum temperature	*Degrees Celsius*	*Min: 3.7* *Max: 34.6* *Mean: 17.4*
Wind speed	*0.1 meter/second*	*Min: 9* *Max: 98* *Mean: 36*
Behaviour: staying at home	*Percentage*	*Min: 0.000* *Max: 0.174* *Mean: 0.078*
Behaviour: traveling	*Percentage*	*Min: 0.115* *Max: 1.000* *Mean: 0.413*
School openings	*I: Normal education. Fully physical.* *II: Mostly physical education / most schools open.* *III: Partially physical, partially non-physical education / Half of schools opened.* *IV: Mostly non-physical education / most schools closed.* *V: All schools closed / only online education.*	40 91 30 44 99
Catering service openings	*I: All catering services normally opened.* *II: All catering services normally opened with distancing measures.* *III: All catering services opened, closing at 12 PM. With distancing measures.* *IV: All catering services opened, closing at 10 PM. With distancing measures.* *V: All catering services closed.*	44 0 121 14 125
Event allowance	*I: All events allowed.* *II: Some events cancelled.* *III: Big events prohibited.* *IV: Events prohibited and gatherings with maximum number of people.* *V: All events and gatherings are prohibited. No maximum.*	40 11 82 53 118
Facemasks	*I: No facemasks.* *II: Facemasks in public transport.* *III: Urge to wear facemasks in all public places.* *IV: Strong urge to wear facemasks in all public places.* *V: Facemasks mandatory in all public places.*	121 119 16 48 0
Number of tests	*Numeric*	*Min: 2,857* *Max: 45,912* *Mean: 17,695*

G. Linear and Interaction regression model

Linear regression model (after including catering service openings)

Predictor	**Estimate**	**Std. Error**	**p-value**
Intercept	1.220	0.177	3.80e-11 ***
Average temperature	-0.013	0.003	8.91e-06 ***
Staying home behaviour	-4.179	0.621	1.06e-10 ***
Traveling behaviour	0.578	0.176	0.001 **
School openings	0.066	0.015	8.41e-06 ***
Catering services	0.021	0.020	0.280
Event allowance	-0.109	0.017	8.99e-10 ***
Facemasks	0.095	0.014	3.79e-11 ***

Polynomial regression model

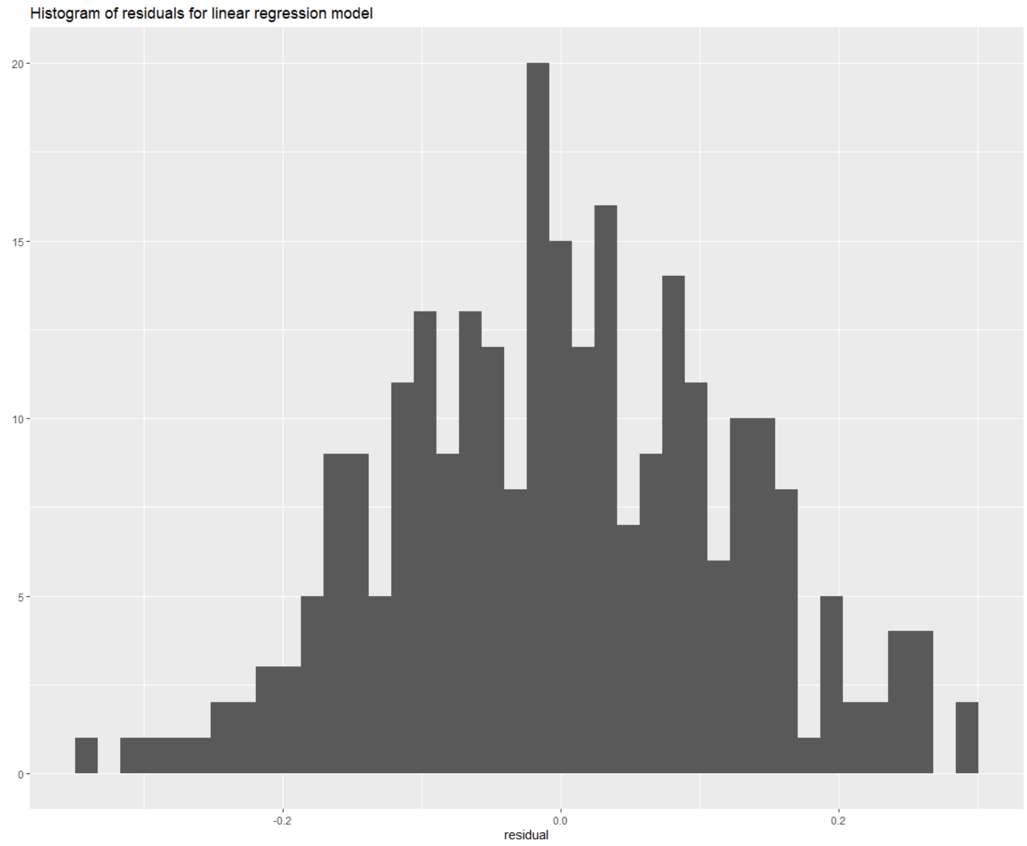
Predictor	**Estimate**	**Std. Error**	**p-value**
(Intercept)	-3.688	2.017	0.069 .
Wind speed	-0.058	0.018	0.001 **
Staying home behaviour	67.41	11.54	1.67e-08 ***
Traveling behaviour	-9.939	3.108	0.002 **
School openings	11.14	3.37	0.001 **
Event allowance	-2.561	0.527	2.11e-06 ***
I(Average temperature^2)	-0.003	0.001	0.003 **
I(Average temperature^3)	0.000	0.000	0.004 **
I(Average temperature^4)	0.000	0.000	0.002 **
I(Wind speed^2)	0.002	0.001	0.001 **
I(Wind speed^3)	0.000	0.000	0.002 **
I(Wind speed^4)	0.000	0.000	0.003 **
I(Staying home behaviour^2)	-1160	198	1.54e-08 ***
I(Staying home behaviour ^3)	7809	1407	7.53e-08 ***
I(Staying home behaviour ^4)	-18180	3472	3.59e-07 ***
I(Staying home behaviour ^2)	41.61	9.67	2.46e-05 ***
I(Staying home behaviour ^3)	-65.33	13.21	1.42e-06 ***
I(Staying home behaviour ^4)	34.57	6.46	1.99e-07 ***
I(School openings^2)	-5.096	1.608	0.002 **
I(School openings^3)	1.000	0.326	0.002 **
I(School openings^4)	-0.071	0.024	0.003 **
I(Catering services openings^2)	0.544	0.106	5.76e-07 ***
I(Catering services openings^3)	-0.105	0.021	6.41e-07 ***
I(Event allowance^3)	0.064	0.014	4.56e-06 ***
I(Facemasks^3)	0.003	0.001	0.003 **

Interaction regression model (after excluding predictors that are not statistically significant)

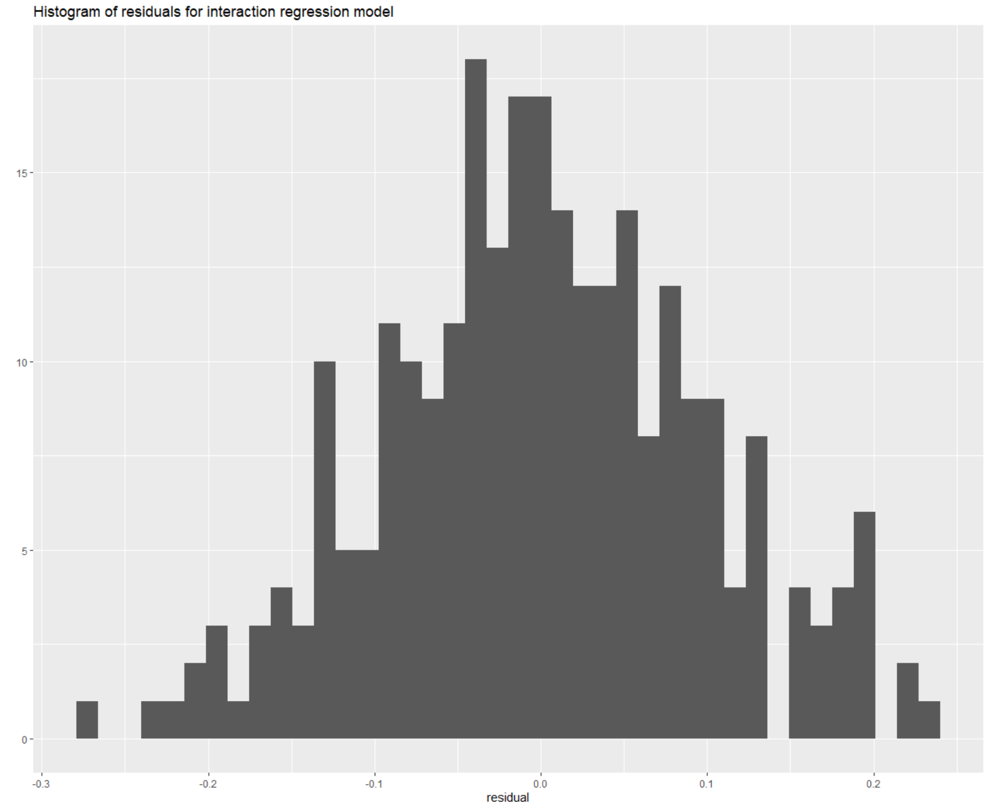
Predictor	**Estimate**	**Std. Error**	**p-value**
Intercept	0.431	0.211	0.042 *
Humidity	0.003	0.001	0.001 **
Wind speed	0.002	0.001	2.79e-05 ***
Staying home behaviour	9.137	1.047	3.58e-16 ***
Traveling behaviour	0.345	0.197	0.080 .
School openings	-0.144	0.030	3.55e-06 ***
Catering openings	-0.216	0.042	6.07e-07 ***
Event allowance	0.064	0.031	0.044 *
Facemasks	0.369	0.083	1.49e-05 ***
*Staying home behaviour*Traveling behaviour*	-20.650	3.784	1.15e-07 ***
*Traveling behaviour*Event allowance*	-0.442	0.108	5.57e-05 ***
*Traveling behaviour*School openings*	0.494	0.088	4.64e-08 ***
*Traveling behaviour*Catering openings*	0.592	0.124	3.04e-06 ***
*Staying home behaviour*Facemasks*	-2.902	0.828	0.001 ***

H. Histogram of residuals for formula Re

Multiple regression model for effective reproduction rate (linear model)

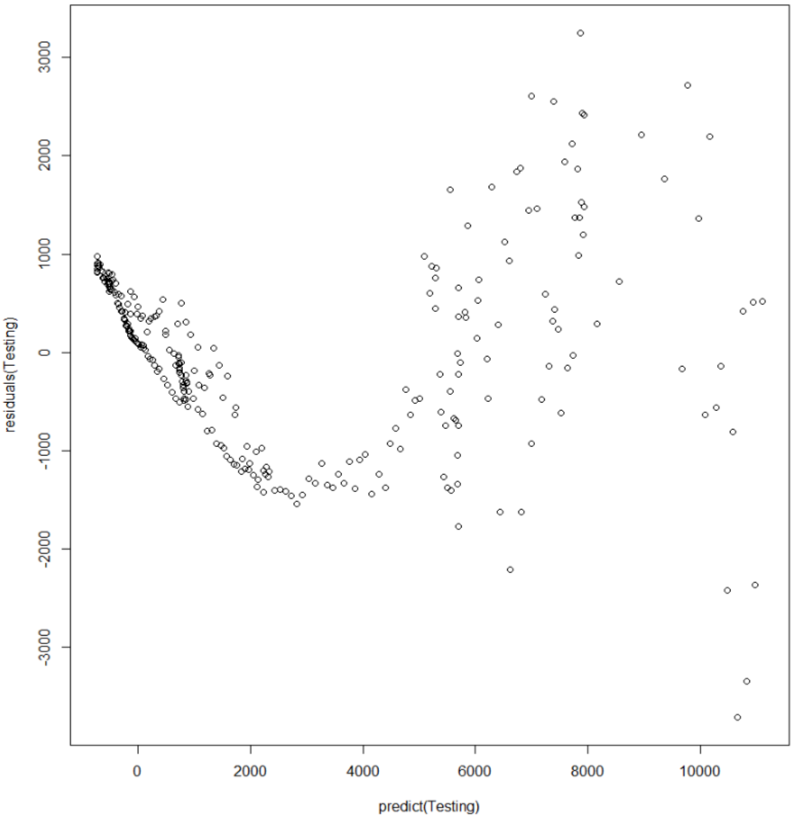


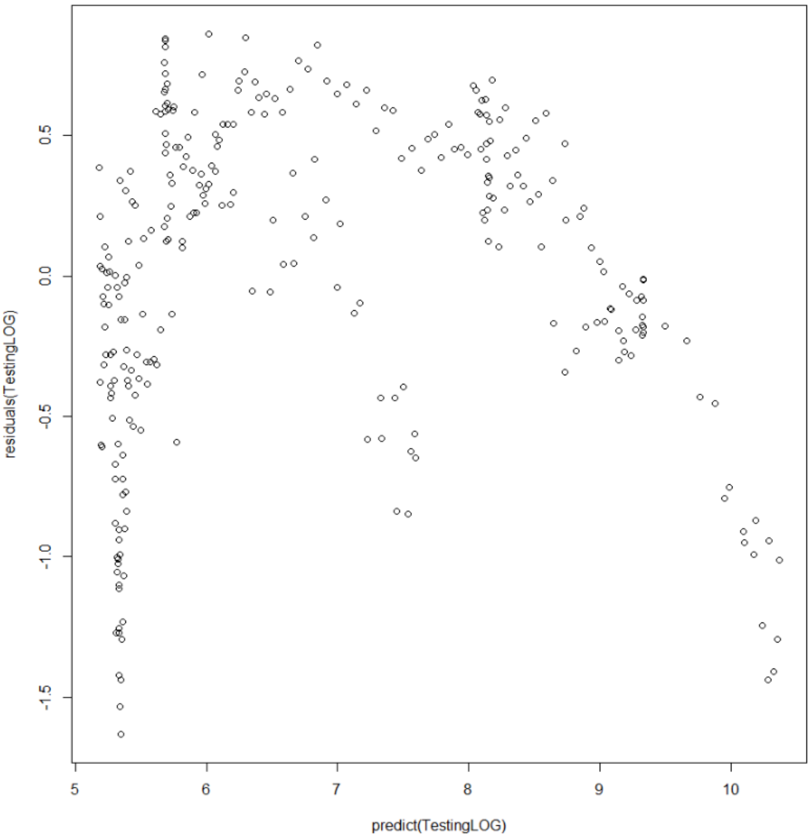
Multiple regression model for effective reproduction rate (linear model with interaction terms)



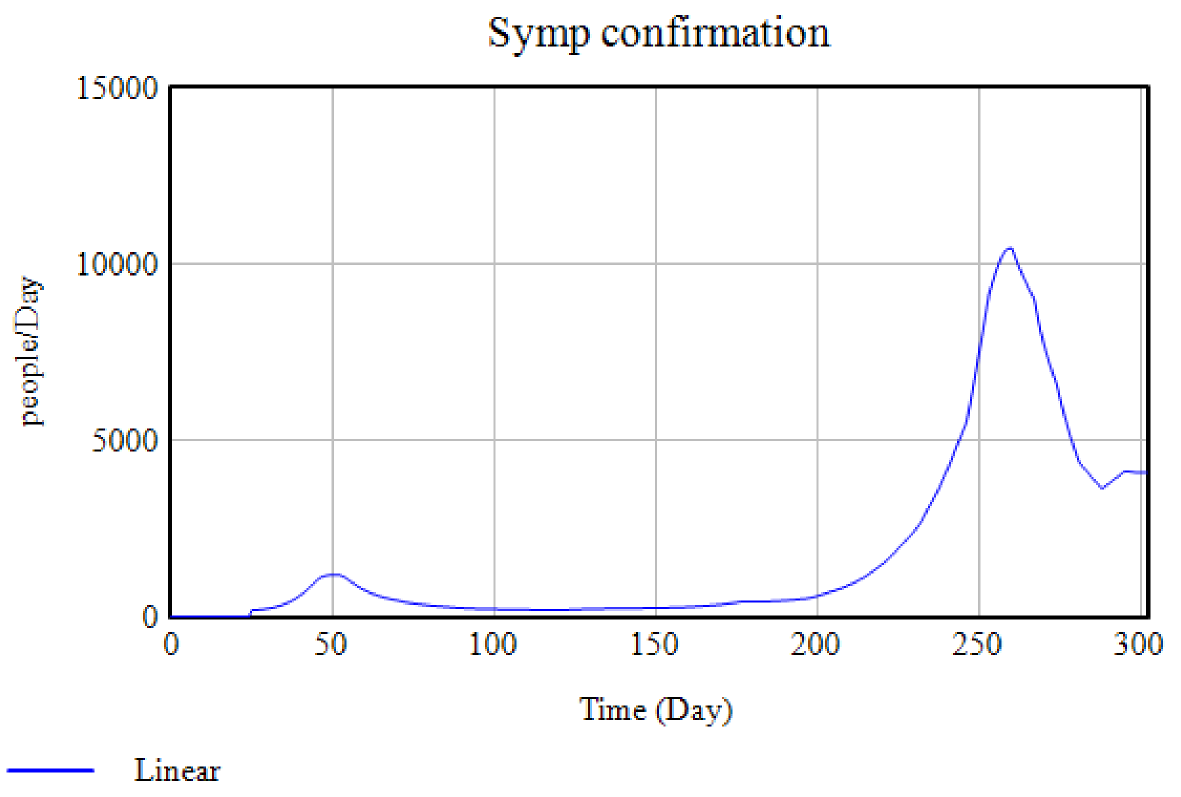
I. Comparison of residual plots for formula number of confirmed cases

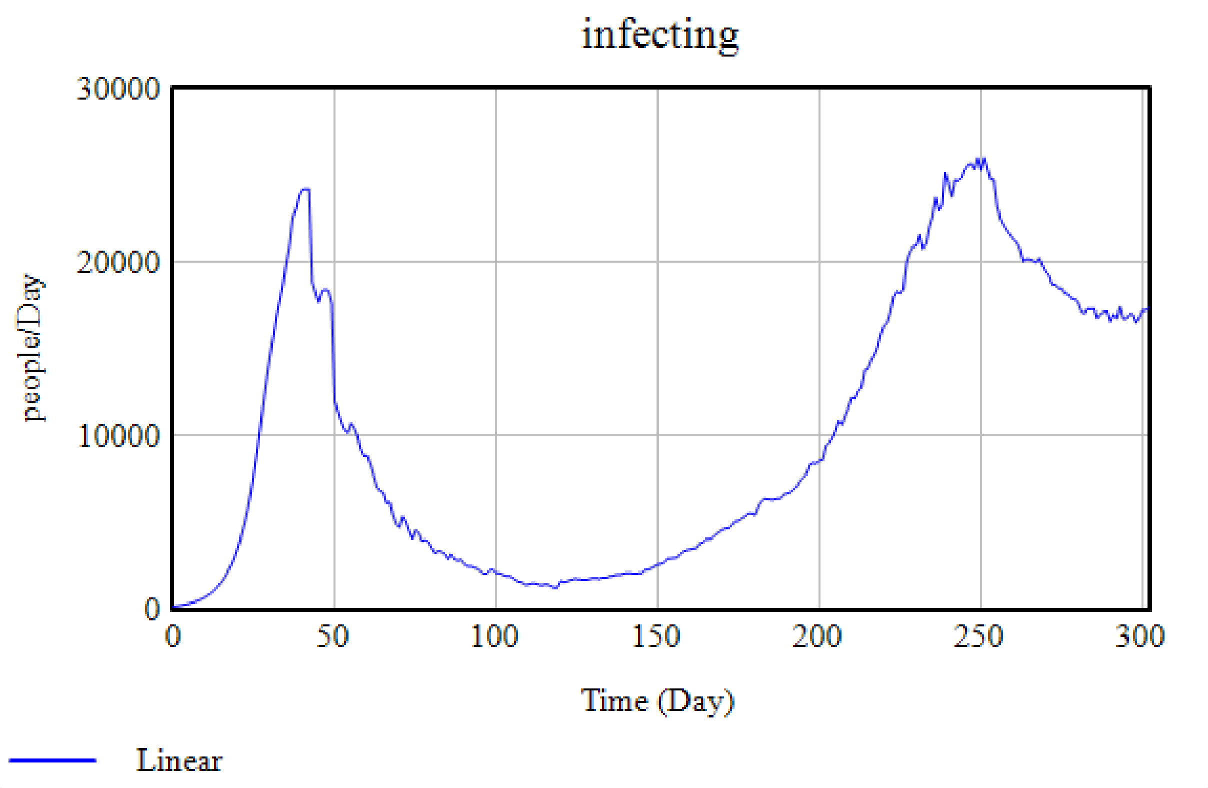
Above residual plot is from the regression model without logarithmic response. Below residual plot is with logaritmic response.

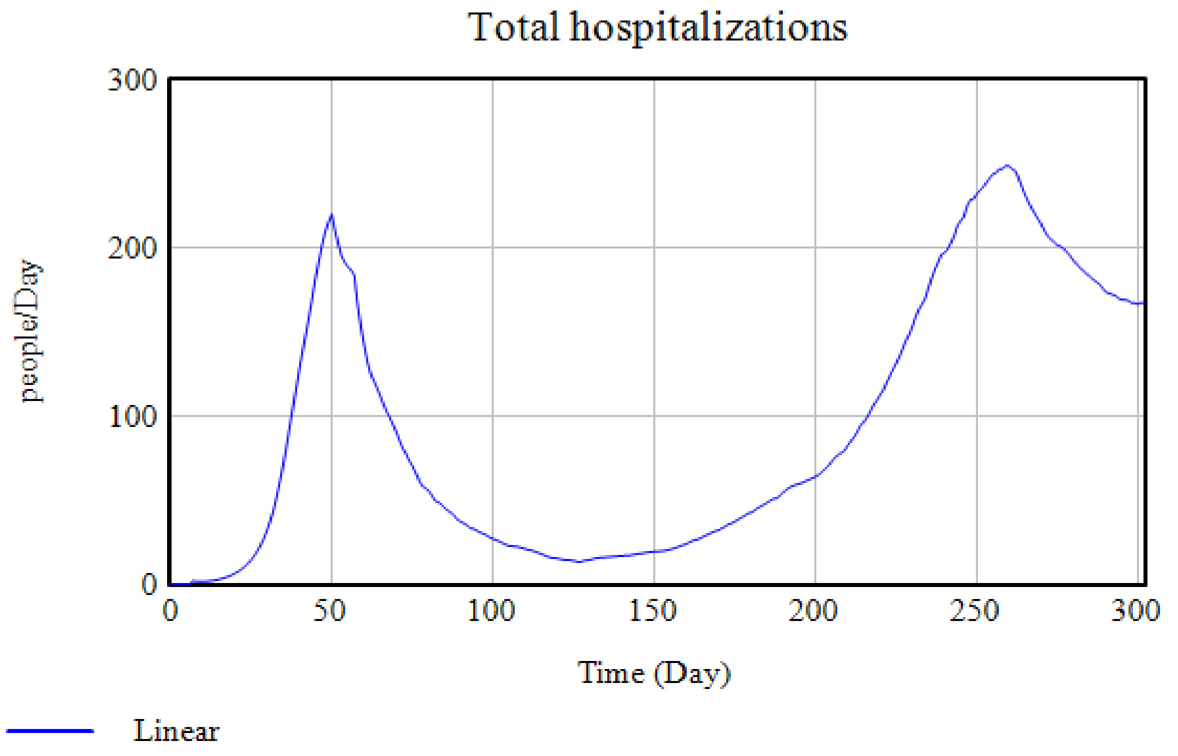


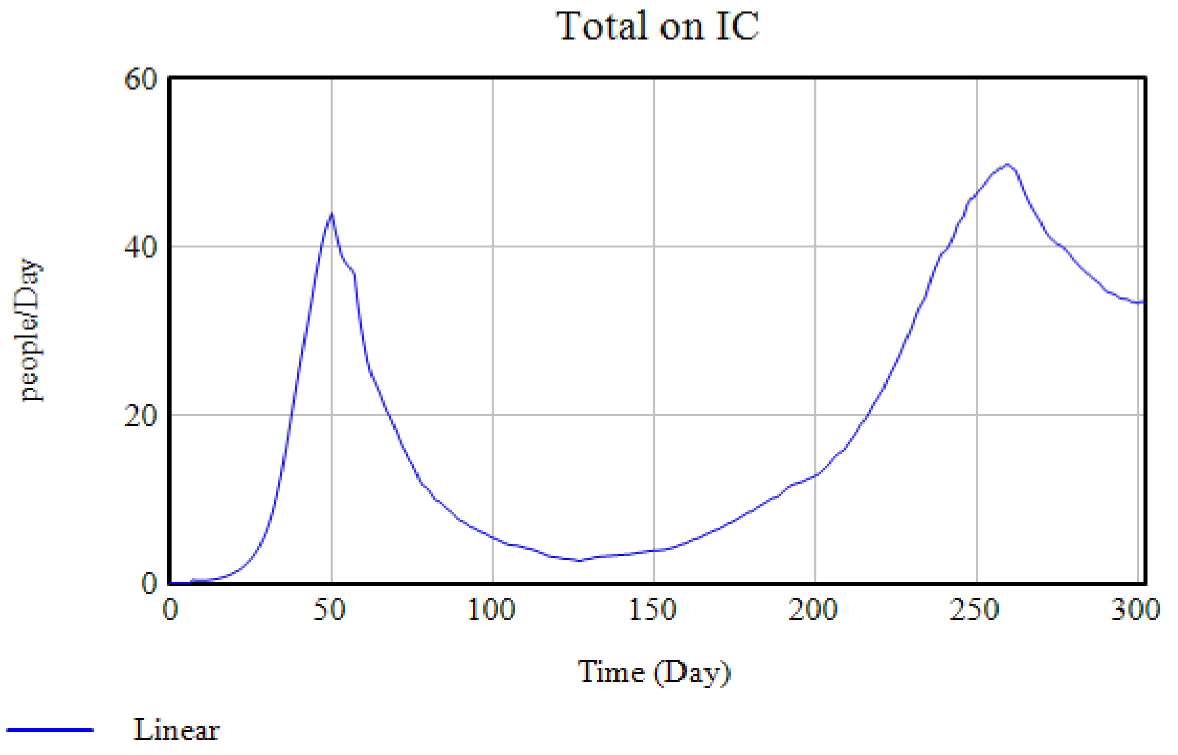


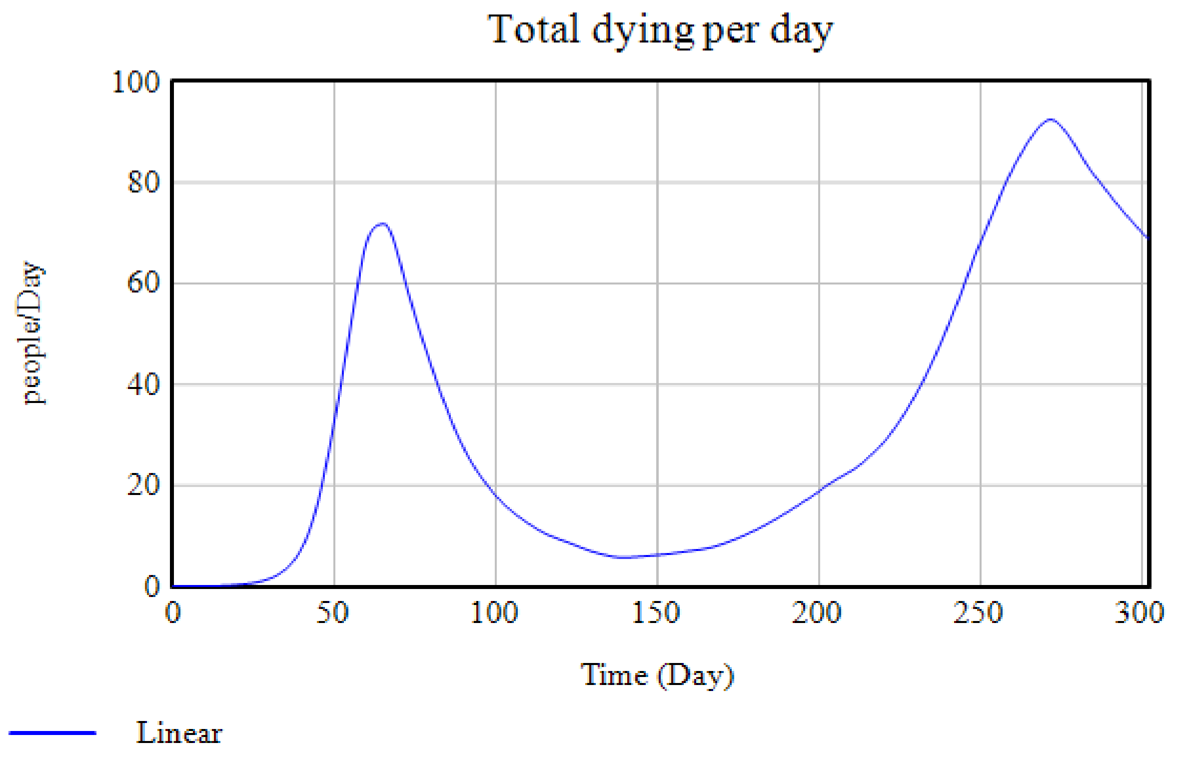
J. Outcomes indicators system dynamics model






Total hospitalizations
300
200
100
0
people/Day
0
50
100
150
200
250
300
Time (Day)
Linear


Total on IC
60
40
20
0
people/Day
0
50
100
150
200
250
300
Time (Day)
Linear


Total dying per day
people/Day
100
80
60
40
20
0
0
50
100
150
200
250
300
Time (Day)
Linear

K. Input values for actual, low and high testing capacity per week

WEEK	DATE	ACTUAL	LOW	HIGH
BEFORE	15/03/2020	2,857	2,857	2,857
12	16/03/2020	3,143	3,143	5,000
13	23/03/2020	3,571	3,571	5,000
14	30/03/2020	3,714	3,714	5,000
15	06/04/2020	5,000	5,000	10,000
16	13/04/2020	5,429	5,000	10,000
17	20/04/2020	5,000	5,000	15,000
18	27/04/2020	3,714	5,000	15,000
19	04/05/2020	3,714	5,000	20,000
20	11/05/2020	4,286	5,000	20,000
21	18/05/2020	3,714	5,000	25,000
22	25/05/2020	4,286	5,000	25,000
23	01/06/2020	6,983	5,000	30,000
24	08/06/2020	8,165	5,000	30,000
25	15/06/2020	8,811	5,000	30,000
26	22/06/2020	8,765	5,000	35,000
27	29/06/2020	9,623	5,000	35,000
28	06/07/2020	10,722	5,000	40,000
29	13/07/2020	12,661	5,000	40,000
30	20/07/2020	15,932	5,000	45,000
31	27/07/2020	14,515	5,000	45,000
32	03/08/2020	14,110	5,000	50,000
33	10/08/2020	15,146	5,000	50,000
34	17/08/2020	19,401	5,000	50,000
35	24/08/2020	22,340	5,000	50,000
36	31/08/2020	25,254	5,000	50,000
37	07/09/2020	27,341	5,000	50,000
38	14/09/2020	27,563	5,000	50,000
39	21/09/2020	29,621	5,000	50,000
40	28/09/2020	31,600	5,000	50,000
41	05/10/2020	39,783	5,000	50,000
42	12/10/2020	44,390	5,000	50,000
43	19/10/2020	45,912	5,000	50,000
44	26/10/2020	41,983	5,000	50,000
45	02/11/2020	34,860	5,000	50,000
46	09/11/2020	32,342	5,000	50,000
47	16/11/2020	36,335	5,000	50,000
48	23/11/2020	37,054	5,000	50,000
49	30/11/2020	45,919	5,000	50,000
50	07/12/2020	63,279		
51	14/12/2020	68,538		
52	21/12/2020	58,655		
53	27/12/2020	49,917		

L. Parameter values per day for mitigation and curbing policies

In the tables below, we provide the date of implementation of new measures from the risk level in the cell to the right (e.g. first date is 15 March 2020, where lockdown measures are implemented).

Mitigation policy

Date	**Risk level**
15/03/2020	Lockdown
30/03/2020	Very serious
18/04/2020	Serious
12/06/2020	Very serious
01/07/2020	Serious
11/09/2020	Very serious
26/09/2020	Serious
28/10/2020	Very serious

Curbing policy 1

Date	**Risk level**
15/03/2020	Lockdown
30/03/2020	Alert
14/04/2020	Alarming
29/04/2020	Very serious
14/05/2020	Serious
29/05/2020	Alarming
13/06/2020	Very serious
28/06/2020	Serious
31/07/2020	Alarming
15/08/2020	Serious
30/08/2020	Alarming
14/09/2020	Serious
29/09/2020	Alarming
28/10/2020	Serious
22/11/2020	Alarming

Curbing policy 2

Date	Risk level
15/03/2020	Lockdown
30/03/2020	Alert
14/04/2020	Lockdown
06/05/2020	Alert
21/05/2020	Lockdown
16/06/2020	Alert
01/07/2020	Lockdown
03/08/2020	Alert

Curbing policy 3

Date	Risk level
15/03/2020	Lockdown
30/03/2020	Alert
13/04/2020	Lockdown
27/04/2020	Alert
11/05/2020	Lockdown
25/05/2020	Alert
08/06/2020	Lockdown
22/06/2020	Alert
06/07/2020	Lockdown
20/07/2020	Alert
03/08/2020	Lockdown
17/08/2020	Alert
31/08/2020	Lockdown
14/09/2020	Alert
28/09/2020	Lockdown
12/10/2020	Alert
26/10/2020	Lockdown
09/11/2020	Alert
23/11/2020	Lockdown

M. Results multivariate sensitivity analysis

Specification of tests

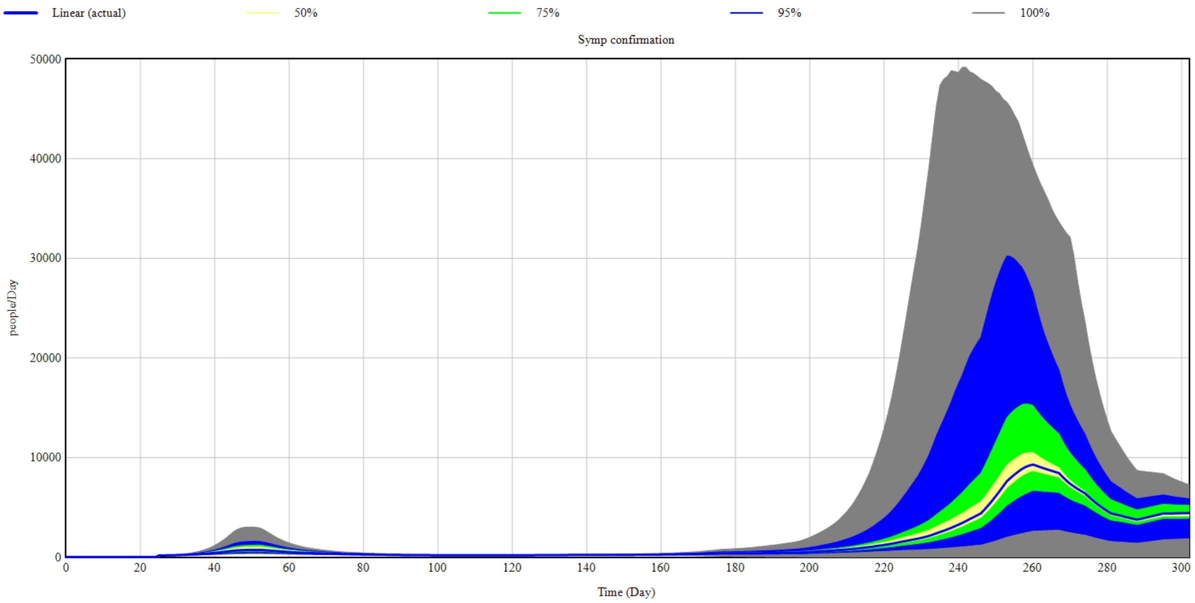
Here we change the basic reproduction rate, initial number of infected cases and quarantine fraction at the same time. Since our model is very sensitive, we let Vensim randomly pick numbers from a normally distributed domain. We do this because we want the program to know which values are likely to fit reality. We provide the distribution of the parameters that are considered in this multivariate sensitivity analysis in the table below.

Parameter	**Current input value**	**Distribution {min; max; mean; standard deviation}**
Basic reproduction rate	3	{2.5; 3.5; 3; 0.15}
Initial infected cases	600	{100; 1,100; 600; 100}
Quarantine fraction	52.5%	{0%; 100%; 50%; 10%}

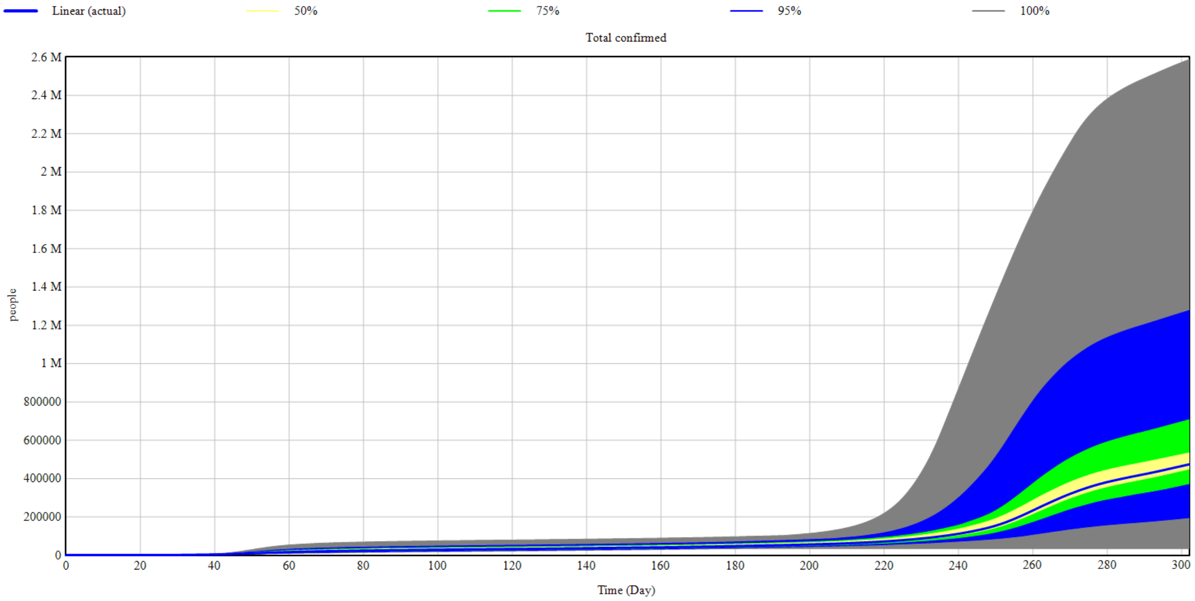
Outcomes of tests

We can see that the total infected cases on 31 November ranges from approximately 500,000 to 5,000,000, and the total confirmed cases ranges from approximately 35,000 to 2,600,000. Especially around the second peak, the confidence interval for the number of confirmed cases per day is wide. Where the first peak ranges from 300 to 3,000, the second peak ranges from 0 to 50,000. This means that when the reproduction rate lies around 2.5, the initial number of infected cases is low, and the quarantine fraction approaches 100%, the spread in the Netherlands would have stopped after the first peak (assuming it would not come back to the Netherlands another time). When the reproduction rate lies around 3.5, the initial number of infected cases is high, and the quarantine fraction approaches 0%, spread could become almost be once as high as in reality. Yet the confidence interval for the number of confirmed cases is different from the confidence of the other indicators. The number of confirmed cases is far more likely to lie around the actual value. Further, the total number of deaths ranges from approximately 1,750 to 17,500. The total number of hospitalizations from 2,900 to 36,000 and the total number of IC occupations from 600 to 7,500.

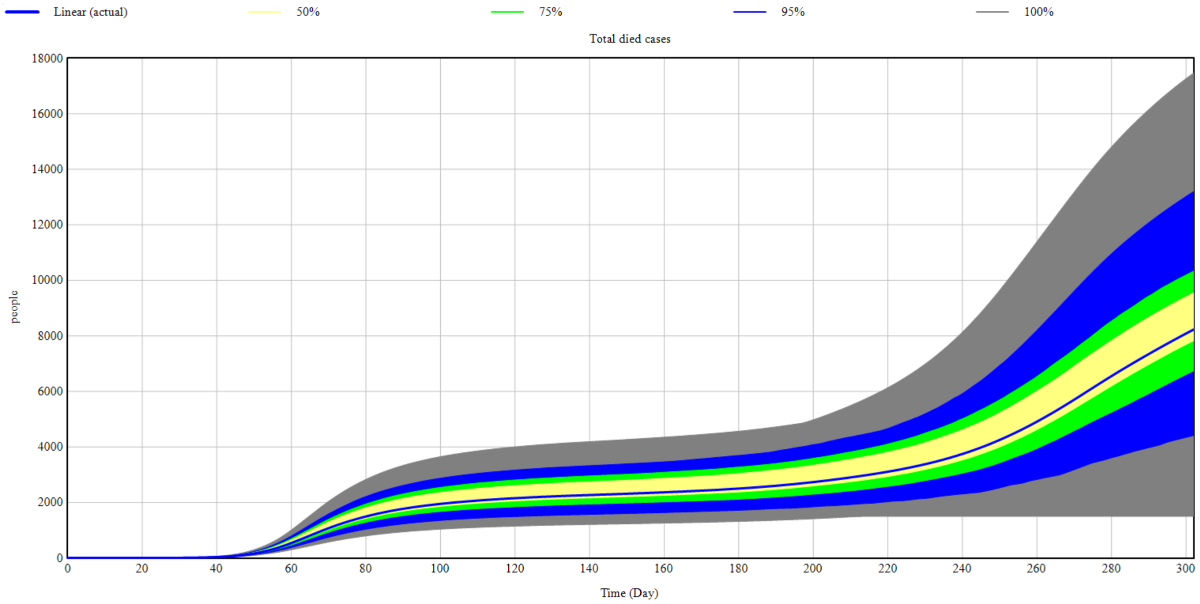
Confirmed cases per day



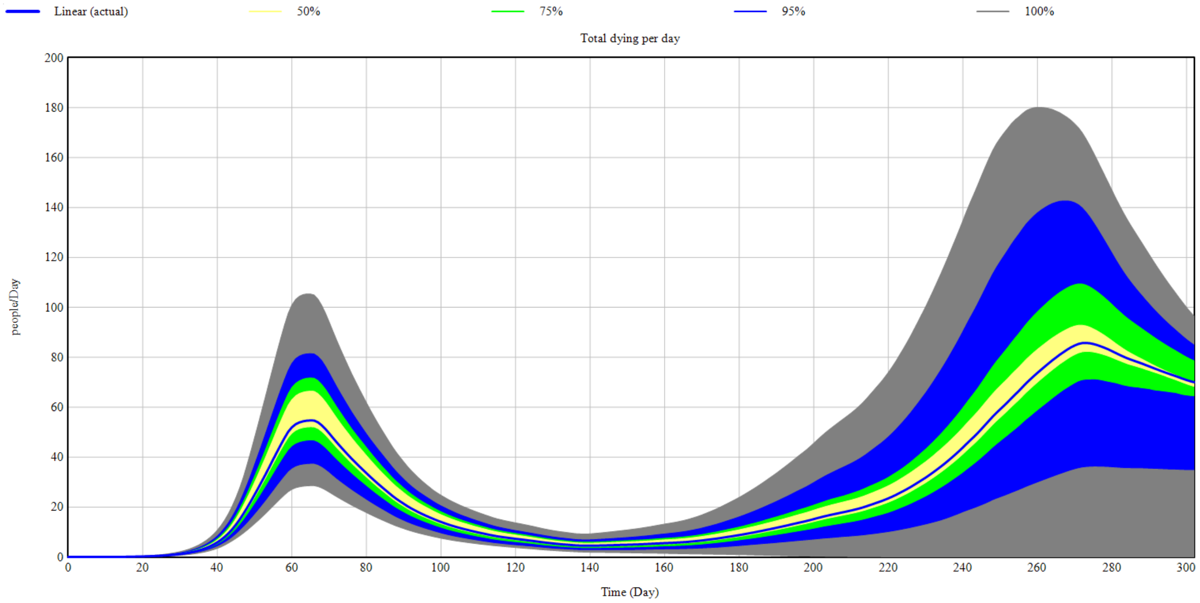
Total confirmed cases



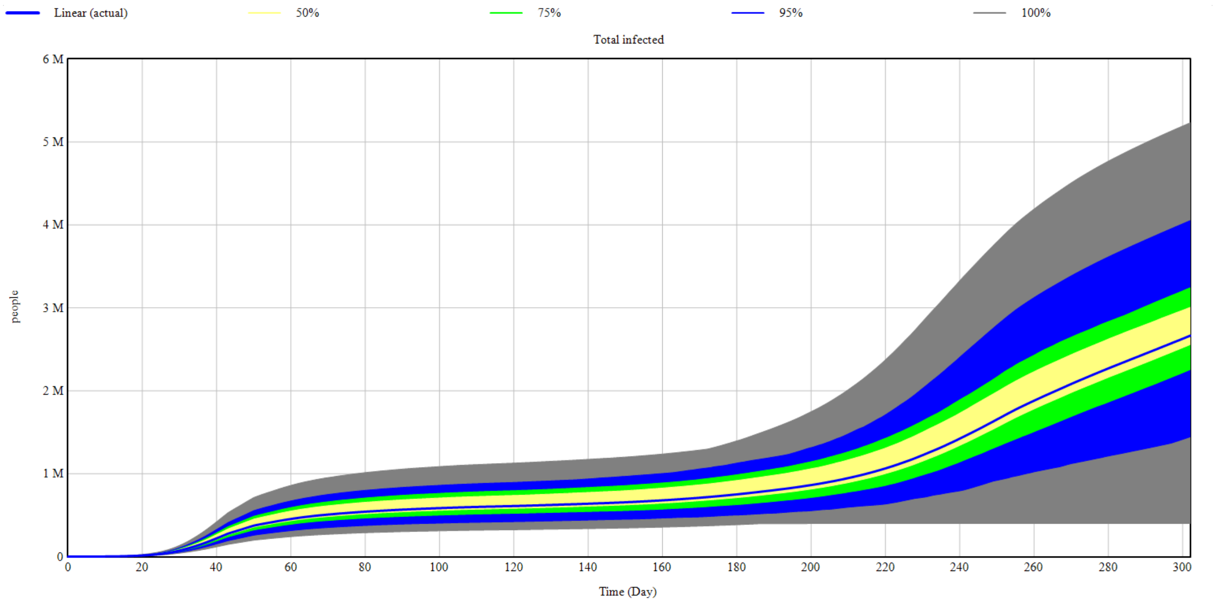
Total died cases



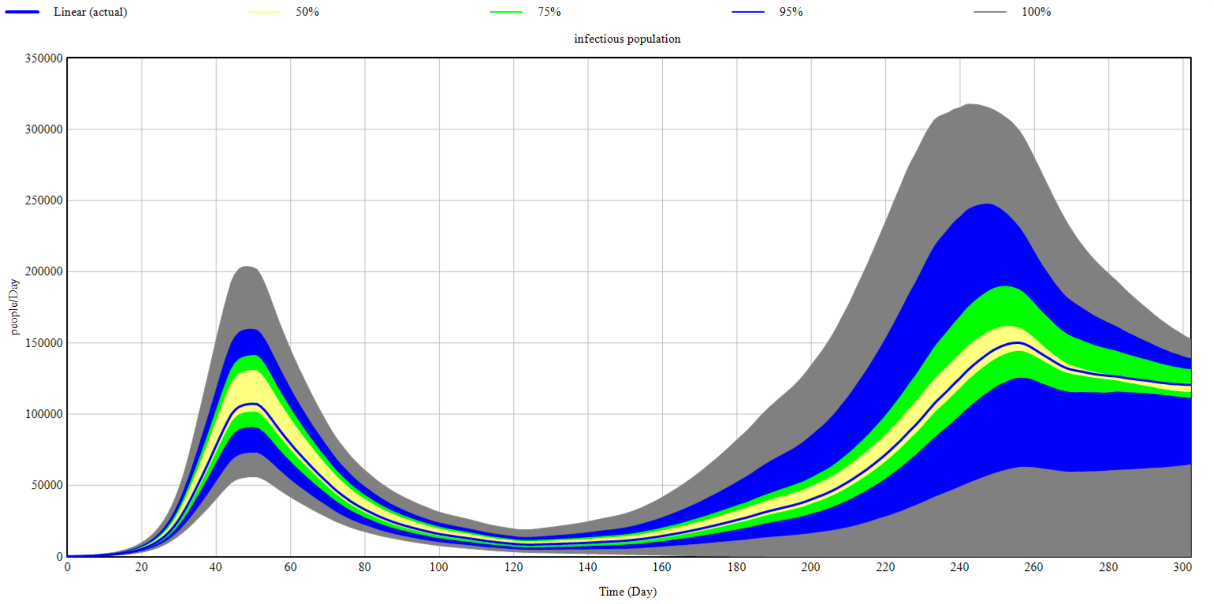
Total dying per day



Total infected cases



Infectious population per day



N. Results univariate sensitivity analysis; Basic reproduction rate and initial cases

Specification of tests

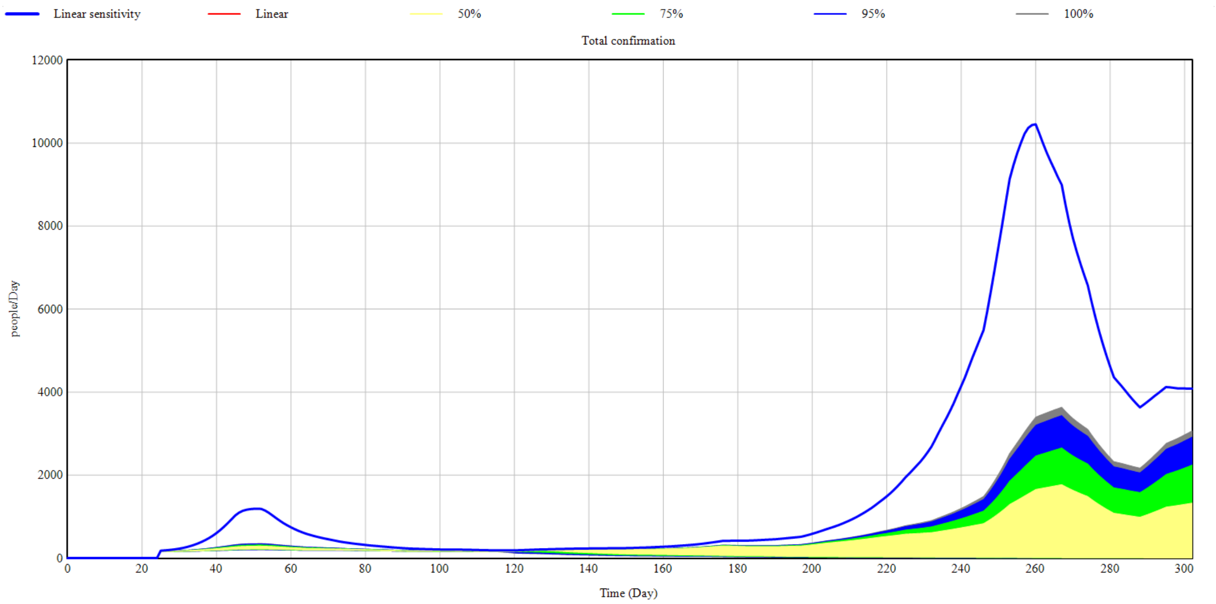
We do four different tests:

1) We vary the basic reproduction rate in different simulations by letting the program pick values from a uniform distribution with a domain between 2 and 2.5.
2) We vary the basic reproduction rate when picking values from a normal distribution with mean 3, minimum 2.5, maximum 3.5 and standard deviation of 0.2.
3) We use the calibrated basic reproduction rate (3) and vary the initial number of infected cases (on 2 February). Here we randomly pick values between 1 and 1000 infected cases.
4) We combine changes of the basic reproduction rate and number of infected cases on 2 February in a multivariate sensitivity analysis. In this multivariate sensitivity analysis we pick values for the basic reproduction rate from a univariate distribution between 2 and 2.5. The number of infected cases in this multivariate analysis ranges between 1 and 2000.

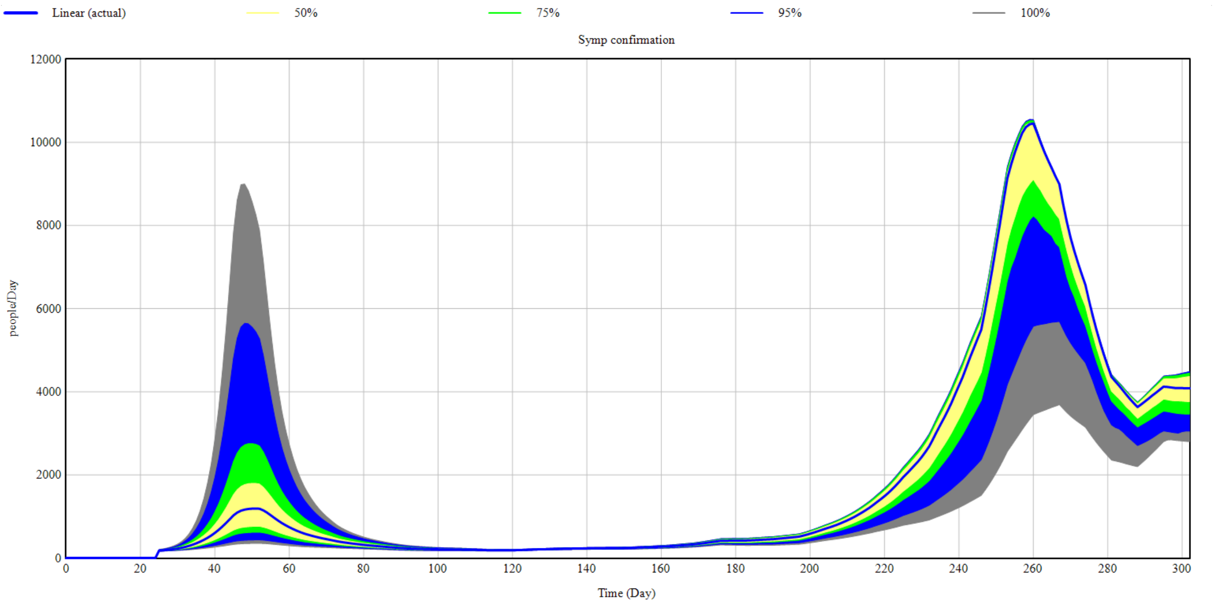
Outcomes of tests

1) We see that when using a basic reproduction rate between 2 and 2.5 instead of 3, the actual number of confirmed cases would by far not be reached. In the first peak, the highest point is approximately 500 cases per day. In the second peak the highest point is approximately 3,750 cases per day.
2) When using a basic reproduction rate between 2.5 and 3.5 with a mean of 3 and a standard deviation of 0.2, the first and the second peak show a large confidence interval. The first peak ranges between approximately 350 and 9,000 confirmed cases per day, where most simulations (50%) provide a value below 2,000. The second peak ranges from approximately 3,750 to 10,500 confirmed cases per day, with most values (50%) above 9,000 at the highest point.
3) When we let the initial infected cases range between 1 and 1,000 with a basic reproduction rate of 3, we obtain a very large confidence interval in the first peak. With the number of confirmed cases per day ranging from approximately 1 to 13,000. Most observations lie below approximately 5,000 at the highest point of the first peak. The confidence interval in the second peak is smaller, ranging from approximately 0 cases per day at the lowest and 10,500 at the highest. Most of the observations in the second peak lie above 6,000 cases at the highest point.
4) When using a basic reproduction rate between 2 and 2.5 and vary the initial number of infected cases between 1 and 2,000, we see that we are able to reach the actual spread. Where the range of the number of confirmed cases in the first peak is small, the range in the second peak is large. In the second peak, 75% of observation lie between 0 and 10,000.

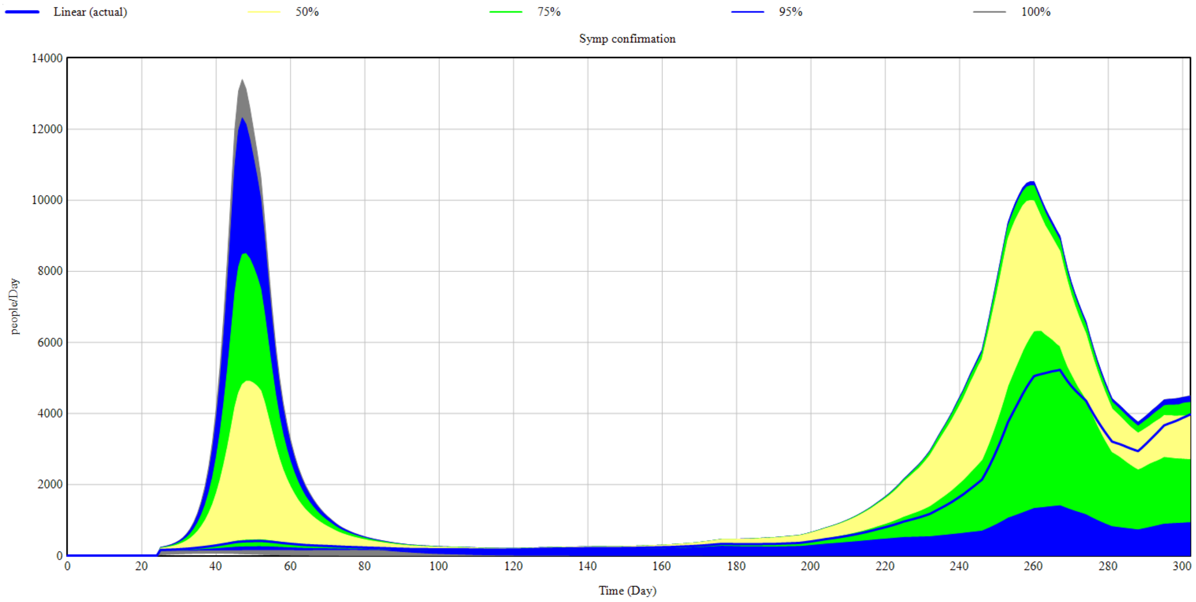
Test 1; Confirmed cases per day



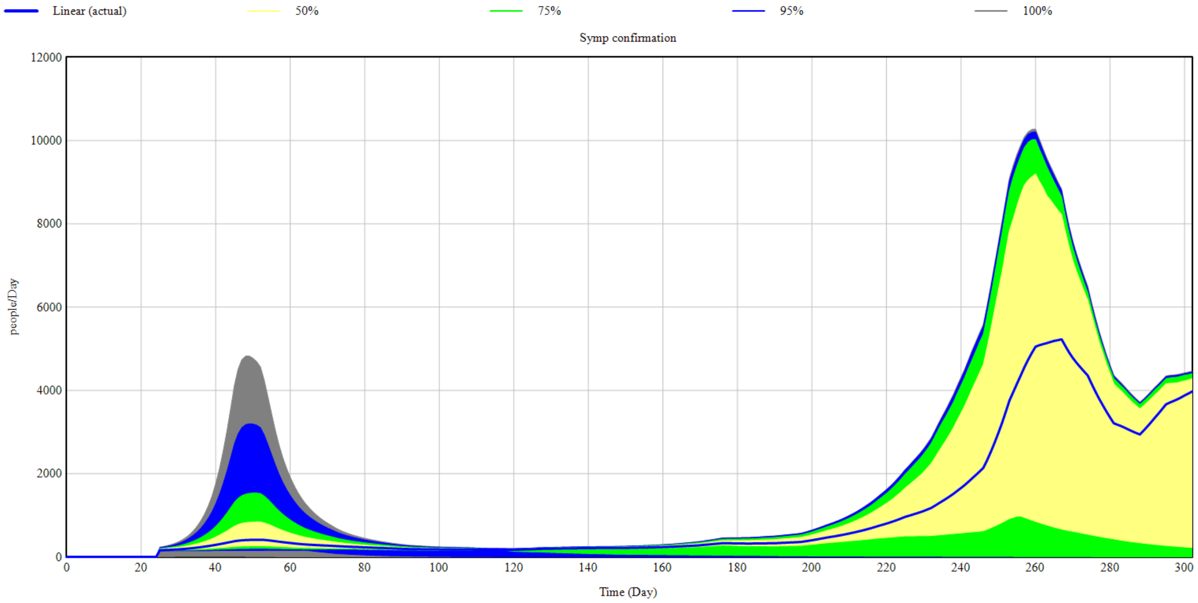
Test 2; Confirmed cases per day



Test 3; Confirmed cases per day



Test 4; Confirmed cases per day



O. Results univariate sensitivity analysis; Quarantine fraction

Specification of tests

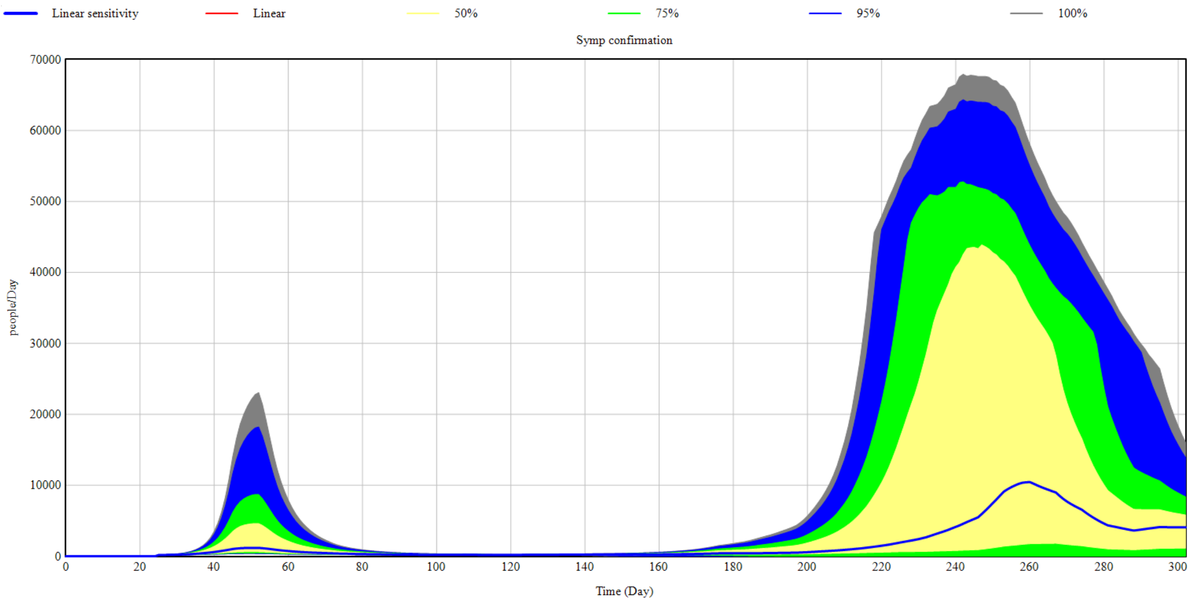
We do two tests:

1) In the first test we vary the quarantine fraction in the actual policy by randomly picking a quarantine fraction from a uniform distribution with a domain from 0% to 100%. We chose this large interval, since evidence for the actual fraction of cases that effectively enters quarantine is lacking. We provide the sensitivity graphs for the number of confirmed cases per day, infectious population, total infected cases and total died cases, to see the effect of these variations.
2) We test what happens to the outcomes of policies when the quarantine fraction would be 80% or 20%. We do this for the feasible policies, being the actual policy, high testing capacity, low testing capacity, curbing (type 1, 2 and 3), and mitigation. Next to a bar graph with the sensitivity results of the total number of infected cases on 31 November, we provide the number of days with strict measures for each policy.

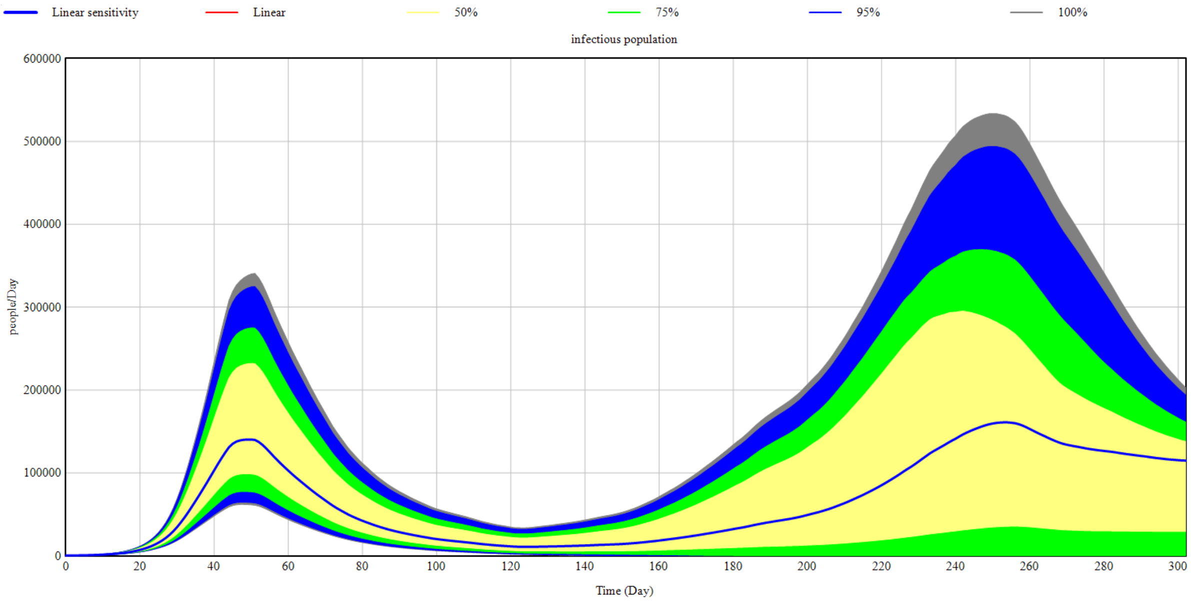
Outcomes of tests

1) We see that when we vary the quarantine fraction between 0% and 100%, this has a major impact on spread. The confirmed cases per day ranges from approximately 300 to 20,000 is the first peak and from zero to almost 68,000 in the second peak. The total infected cases ranges from approximately 500,000 to 7,500,000 and the total deaths ranges from approximately 2,000 to 26,000. The blue line in the graphs indicates the actual spread, with a quarantine fraction of 52.5%.
2) In the bar graph (last figure), we provide results of the sensitivity tests for all feasible policies with a quarantine fraction of 52.5%, 20% and 80% respectively. We observe that all policies show the lowest total number of infected cases with a quarantine fraction of 80%. Based on the graph, it seems to differ per policy how effective quarantine is. Especially for the mitigation, curbing type 1 and curbing type 2 policy, the effect on spread of a smaller or larger quarantine fraction seems to be smaller. However, the number of days with strict measures changes for these policies with a different quarantine fraction. For curbing type 3, low testing capacity, high testing capacity, and the actual policy, the number of days with certain measures remains the same. When we apply the mitigation policy, all days (260) have serious, very serious or lockdown measures. Mitigation with a quarantine fraction of 80% has one short lockdown at the beginning and only serious measures thereafter. Mitigation with a quarantine fraction of 20% has almost an equal number of days with serious measures as with very serious measures. The policy with a high testing capacity and a quarantine fraction of 80% shows the lowest total number of infected cases, and the actual policy the lowest total number of infected cases thereafter. Yet the number of days with strict measures in these two policies is above 200. With curbing type 3, the spread is lower than curbing type 1 and curbing type 2, yet the number of days with strict measures is highest.

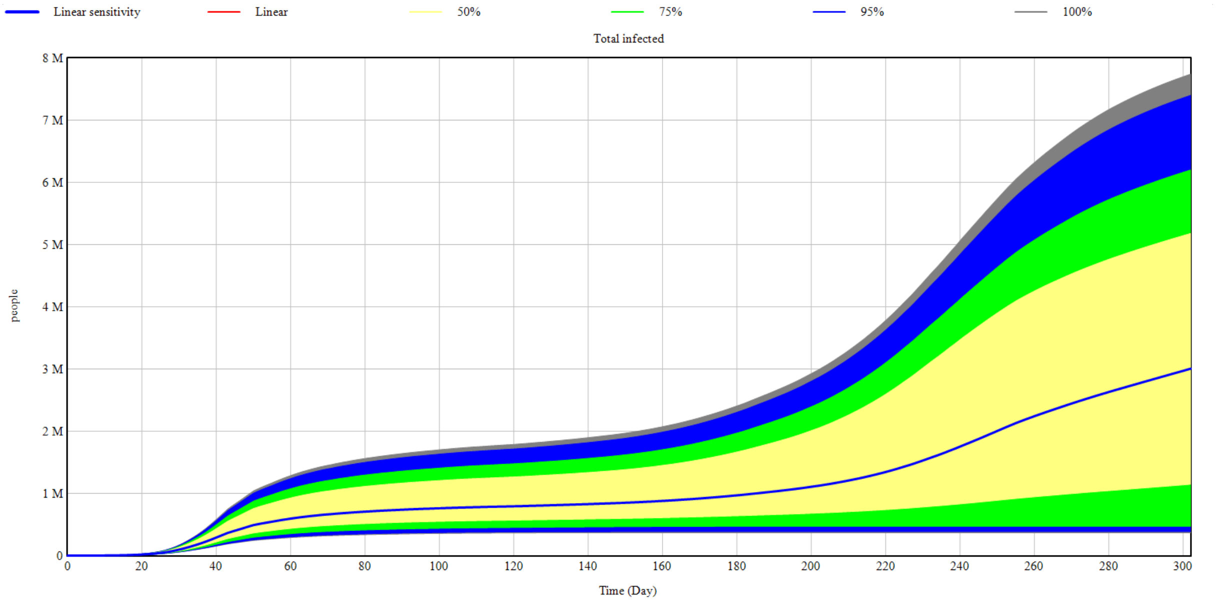
Test 1; Confirmed cases per day



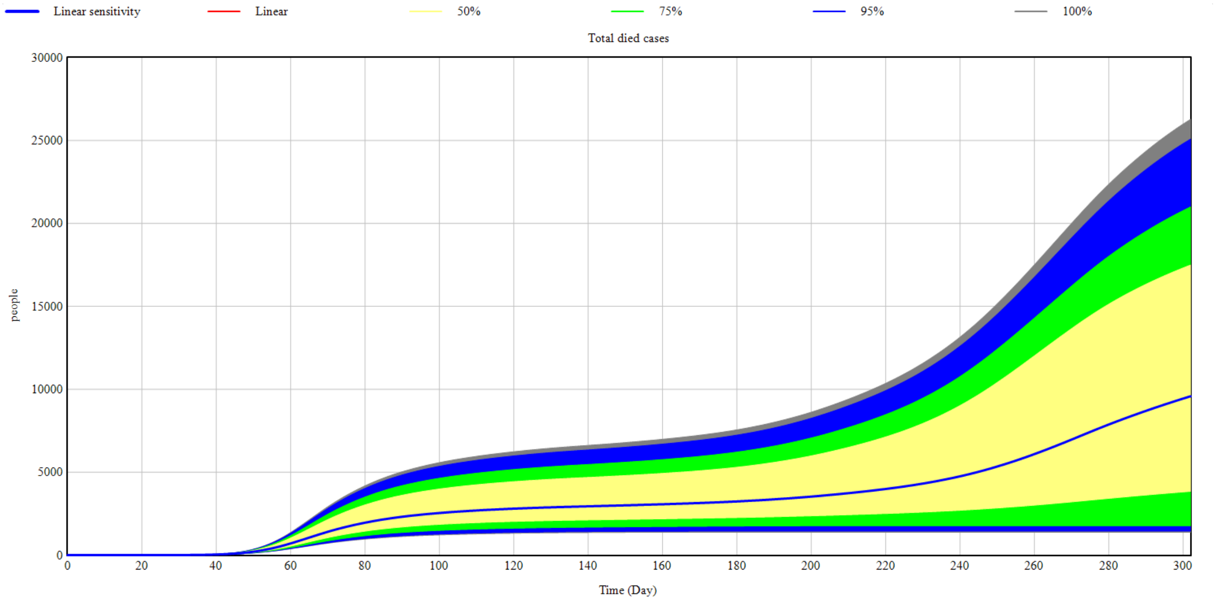
Test 1; Infectious population per day



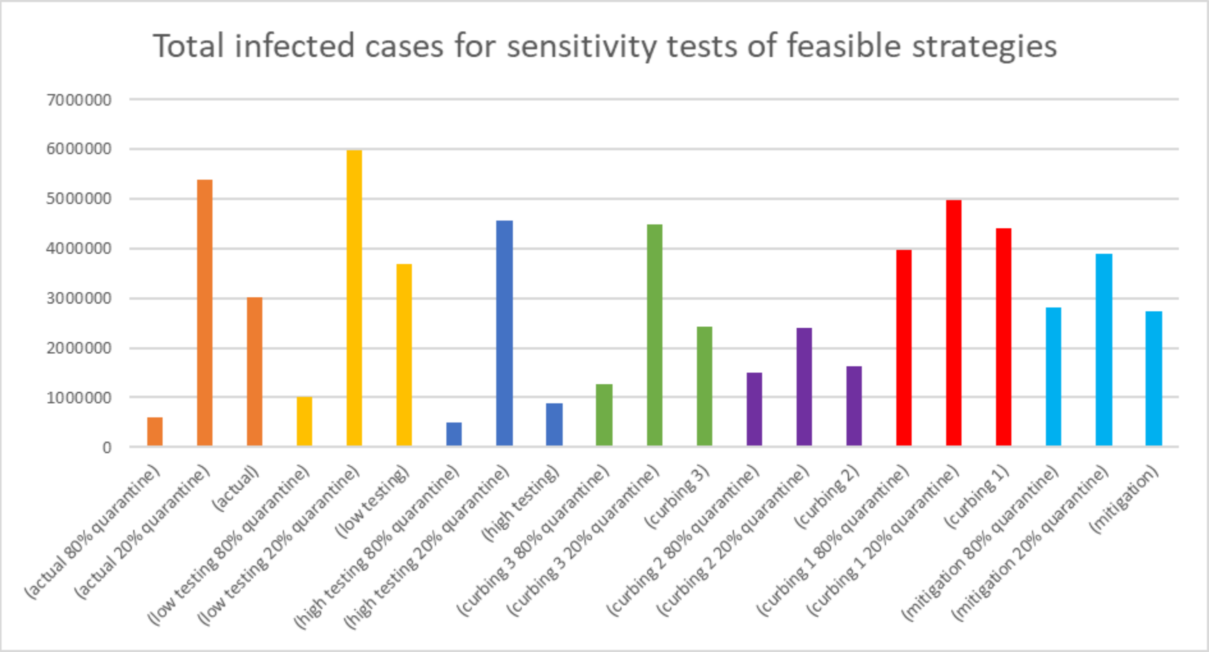
Test 1; Total infected cases



Test 1; Total died cases



Test 2; Total infected cases



Test 2; Days with strict measures

Policy	**Days with strict measures**
Actual	+/- 200
Mitigation (20%)	260
Mitigation (52.5%)	260
Mitigation (80%)	260
Curbing 1 (20%)	200
Curbing 1 (52.5%)	158
Curbing 1 (80%)	99
Curbing 2 (20%)	157
Curbing 2 (52.5%)	96
Curbing 2 (80%)	83
Curbing 3 (20%)	140
Curbing 3 (52.5%)	140
Curbing 3 (80%)	140

P. Results univariate sensitivity analysis; Asymptomatic fraction (infectious period)

Specification of tests

We determine the effect of a change in asymptomatic fraction by changing the duration of the infectious period. We do it this way, since asymptomatic cases have a shorter infection duration and consequently shorter infectious period. When rewriting the formula for the infectious rate on page 45, we can determine the infectious period $\frac{1}{\gamma}$, with v_s the symptomatic fraction and v_a the asymptomatic fraction. This gives the following formula for the duration of the infectious period:

$$\frac{1}{\gamma} = v_s \frac{1}{\gamma_s} + v_a \frac{1}{\gamma_a}$$

In the literature review we found the infectious period of asymptomatic cases to be approximately 5.5 days and the infectious period of symptomatic cases approximately 9 days. Considering these values, the infectious period in our system dynamics model should have been 7.4 days with 40% of the cases being asymptomatic.

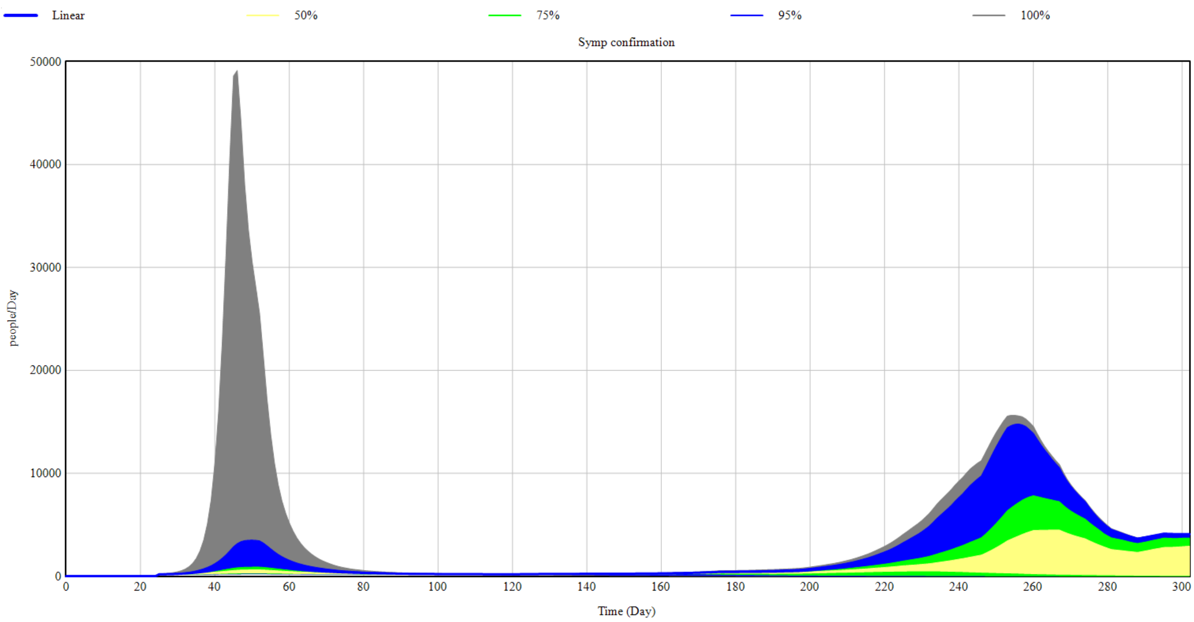
We do two sensitivity tests:

1) We randomly pick values for the infectious duration from a normally distributed domain with mean 7.4 days, standard deviation of 1 day, a minimum of 4.4 days, and a maximum of 10.4 days.
2) We consider a mean of 6 days, standard deviation of 0.2 days, a minimum 5 days and maximum 7 days.

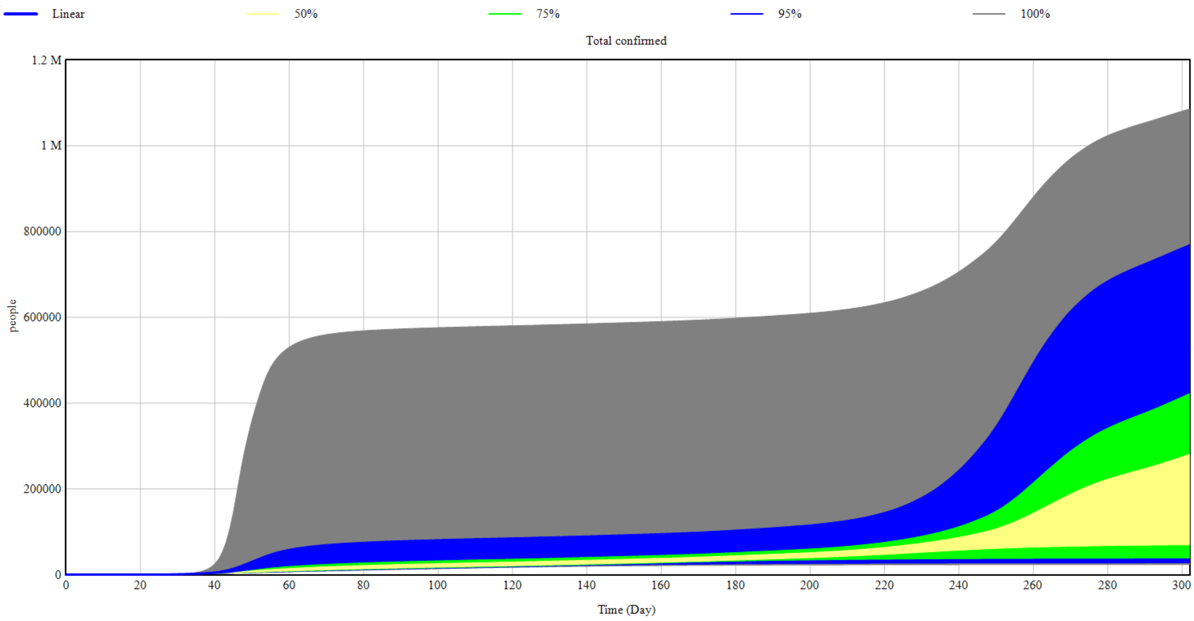
Outcomes of tests

1) We see that confidence interval is large for all indicators. What is remarkable is that the first peak of number of confirmed cases per day reaches almost 50,000, while the second peak reaches maximally 15,000. Yet more than 95% of the simulations has a highest point of the first peak that lies below 3,500. This means that the infectious period is especially sensitive during the first peak. The total number of infected cases for this test ranges between approximately 100,000 and 5,300,000, the total number of confirmed cases ranges between approximately 24,000 and 1,080,000, and the total number of deaths between approximately 400 and 18,400.
2) We observe that the ranges are much smaller but still quite large. The total infected cases ranges from approximately 2,000,000 to 4,000,000 and the total number of confirmed cases from approximately 300,000 to 750,000. Concluding, the infectious period and thus the fraction of asymptomatic cases is a sensitive parameter in our model.

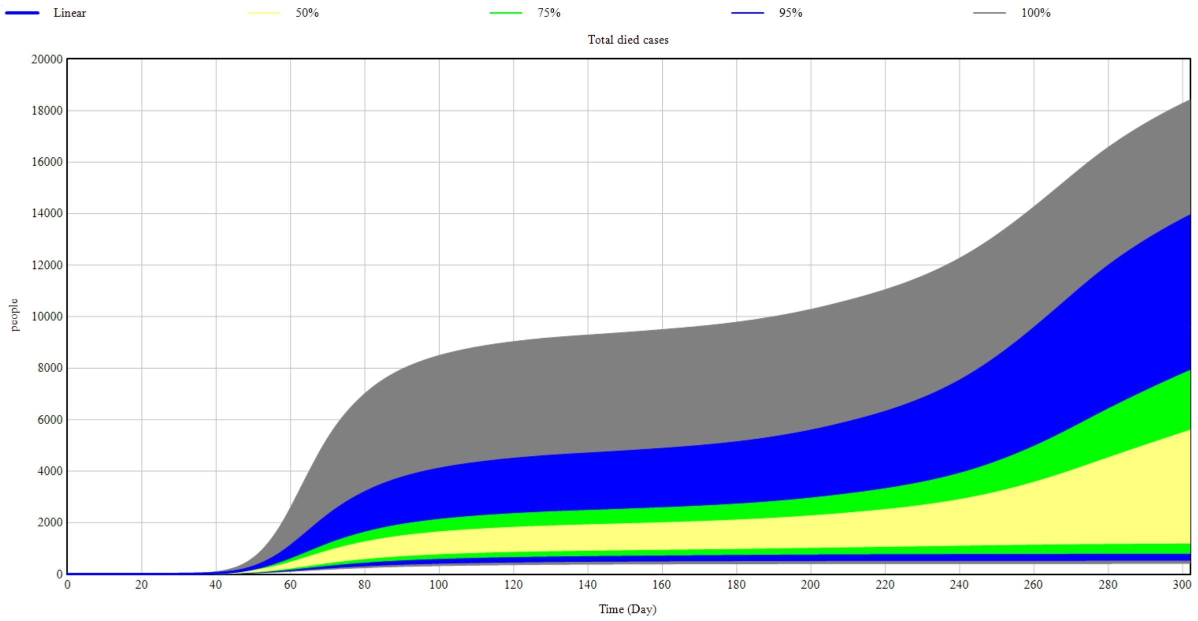
Test 1; Confirmed cases per day



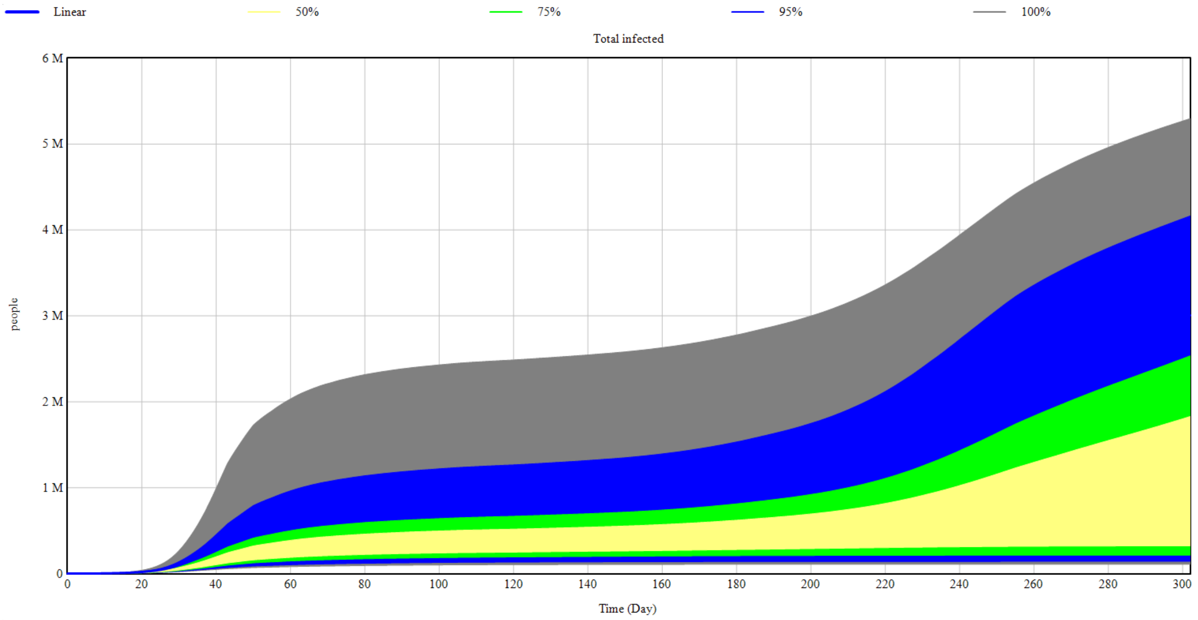
Test 1; Total confirmed cases



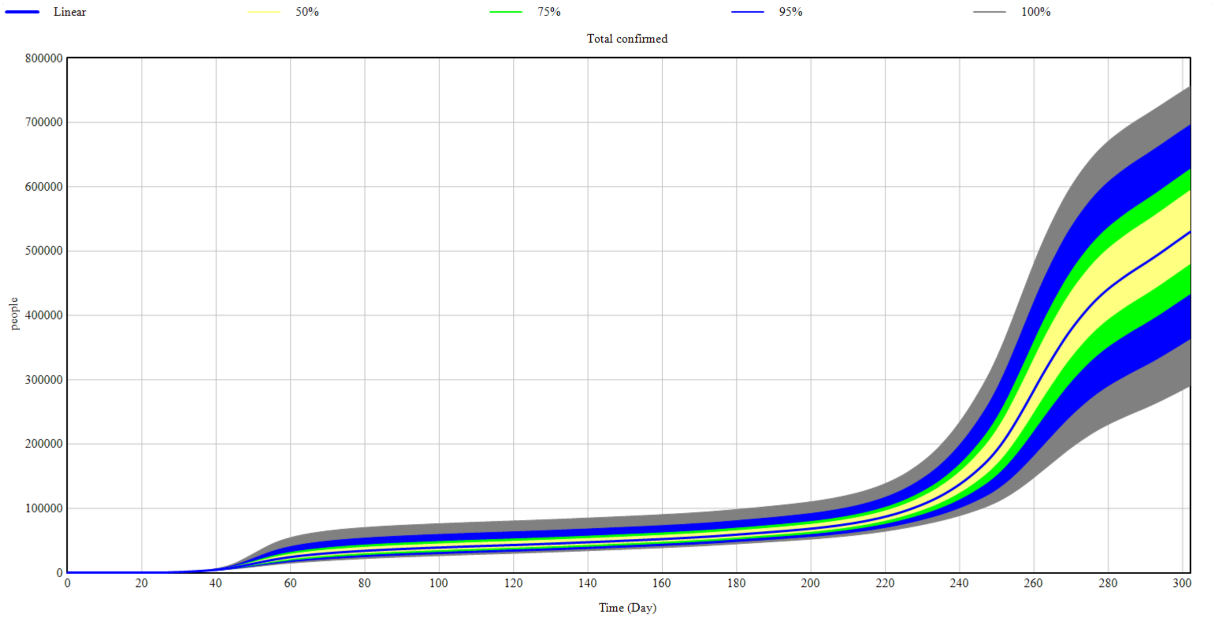
Test 1; Total died cases



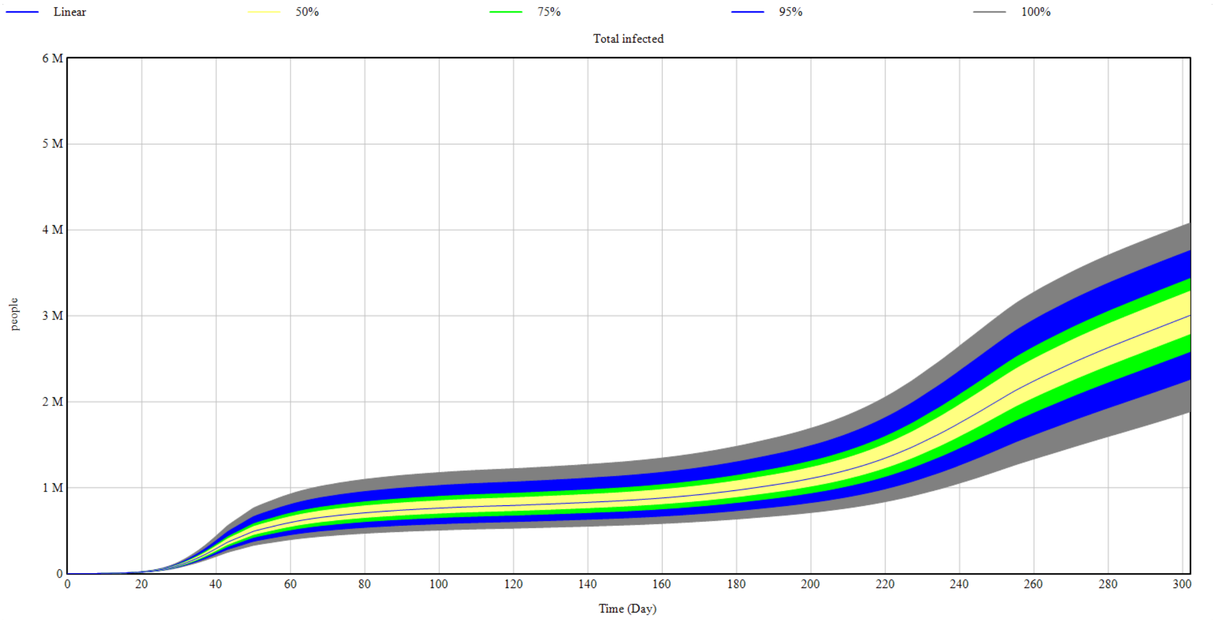
Test 1; Total infected cases



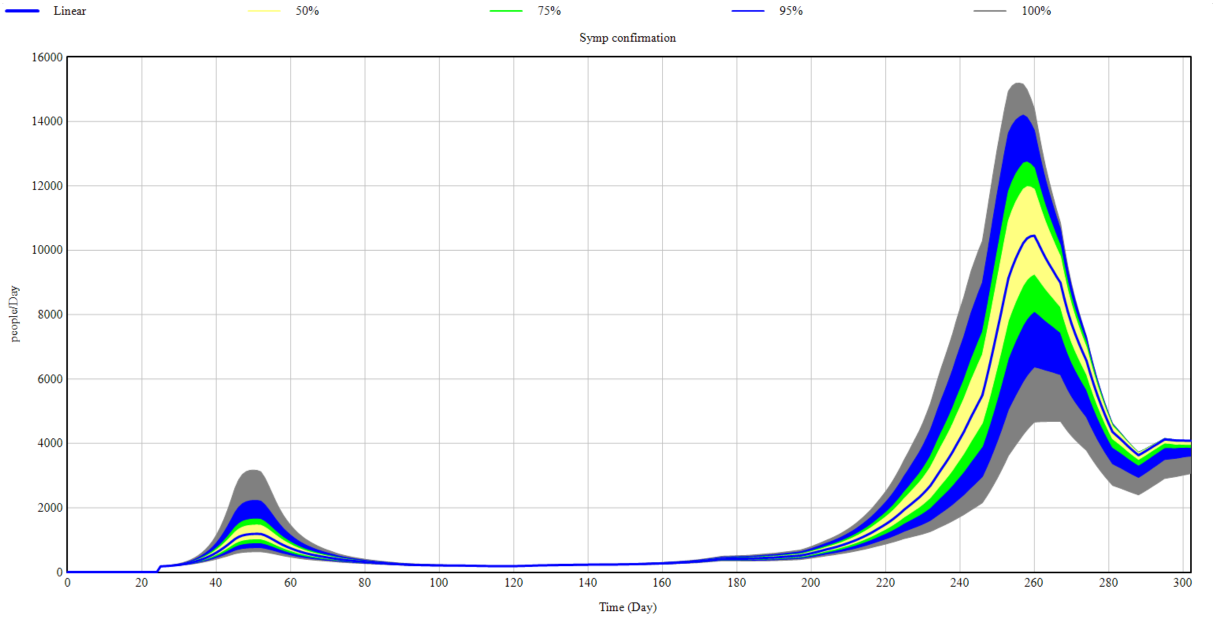
Test 2; Total confirmed cases



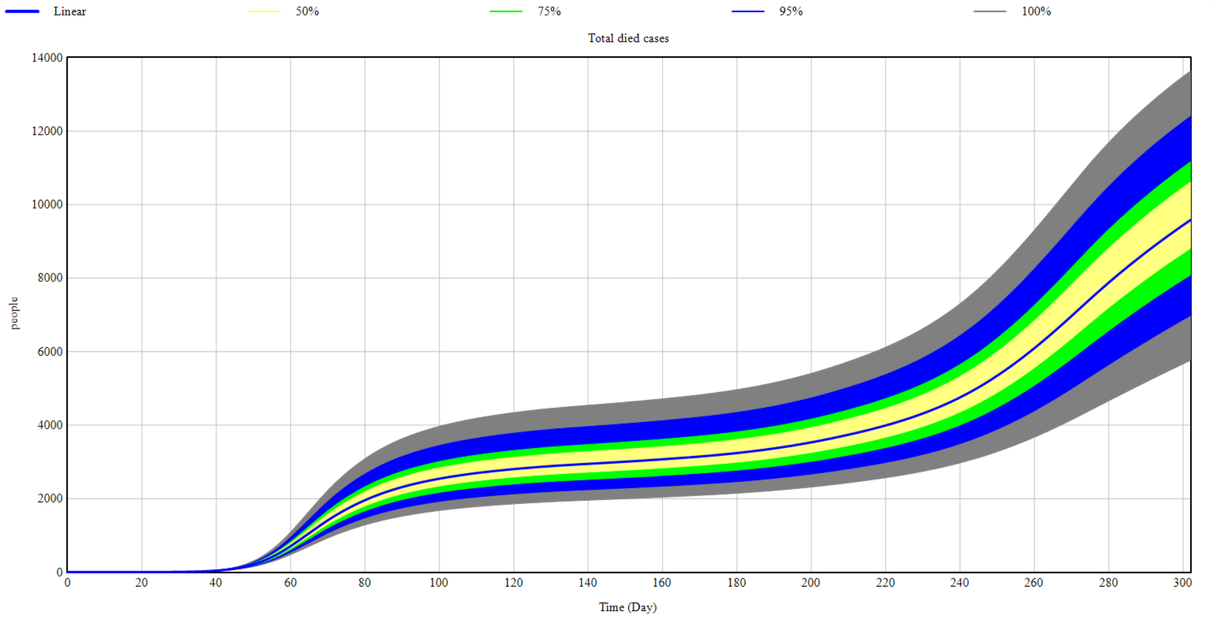
Test 2; Total infected cases



Test 2; Confirmed cases per day



Test 2; Total died cases



Q. R-code

Below, we provide in **bold** the references to the report.

```
# LOAD LIBRARIES -------------------------------------------------
library(ISLR)
library(leaps)
library(dplyr)
library(forcats)
library(psych)
library(caret)
library(randomForestSRC)
library(MASS)
library(rms)
library(ROCR)
library(pROC)
library(DBI)
library(RPostgreSQL)
library(readr)
library(tidyr)
library(lubridate)
library(lmtest)
library(ggplot2)
library(boot)
library(bootStepAIC)
library(R0)
library(car)

# DATA PREPERATION -------------------------------------------------
DATA <- read_delim(file = "C:/Users/cjten/Desktop/NewestCSV.csv",
                   delim = ",", locale = locale(encoding = "ISO-8859-1"),
                   col_names = TRUE, col_types = NULL)

TestingDATA <- read_delim(file = "C:/Users/cjten/Desktop/Testing.csv",
                   delim = ",", locale = locale(encoding = "ISO-8859-1"),
                   col_names = TRUE, col_types = NULL)

R_DATA <- read_delim(file = "C:/Users/cjten/Desktop/incidence1512.csv",
                   delim = ",", locale = locale(encoding = "ISO-8859-1"),
                   col_names = TRUE, col_types = NULL)

R2DATA <- within(DATA,{
  Date <- NULL
  Percentage_pos_tests <- NULL
  Positive_tests_per_day <- NULL
  R_RIVM <- NULL
  Confirmed_cases <- NULL
  Confirmed_infected <- NULL
  Hospitalized_infected <- NULL
  Estimated_infections <- NULL
  Hospitalized_RIVM <- NULL
  IC_occupations <- NULL
  Deaths_RIVM <- NULL
  Deaths_infected <- NULL
})

incidence <- within(R_DATA,{
  Re_infections <- NULL
  Re_hospitalizations <- NULL
  Re_RIVM <- NULL
  Confirmed_RIVM <- NULL
  Hospitalized_RIVM <- NULL
})
```

```
# DESCRIPTIVE STATISTICS
FrequentieTabel1 <- table(R2DATA$Facemasks)
FrequentieTabel1

FrequentieTabel2 <- table(R2DATA $Schools)
FrequentieTabel2

FrequentieTabel3 <- table(R2DATA $Catering)
FrequentieTabel3

FrequentieTabel4 <- table(R2DATA $Events)
FrequentieTabel4


# REMOVING ROWS WITH LACKING DATA ------------
dim(TestingDATA)
TestingDATA <- na.omit(TestingDATA)
dim(TestingDATA) #

dim(R2DATA)
R2DATA <- na.omit(R2DATA)
dim(R2DATA) #

dim(incidence)
incidence <- na.omit(incidence)
dim(incidence)

## NORMALIZATION EFFECT ---------------
## Normalize the input variables
normalize <- function(x){
  return((x-min(x))/(max(x)-min(x)))}
TransformedDataR <- as.data.frame(lapply(R2DATA[c(4,5,6,7,8,9,10,11,12,13,14)], normalize))
InputDataR <- cbind(R2DATA[c(1,2,3)],TransformedDataR[c(1,2,3,4,5,6,7,8,9,10,11)])


# ESTIMATING Re CALCULATED (Section 4.2.1) -----------------
# Size of mean generation interval
mGT <- generation.time("gamma", c(3.99,2.95))

# Develop estimates for Re
estR <- est.R0.TD(incidence$Confirmed_infected,mGT,begin = 1, end = 300, nsim = 1000)
estR2 <- est.R0.TD(incidence$Hospitalized_infected,mGT,begin = 1, end = 300, nsim = 1000)

# Plot calculated reproduction rates (Appendix E)
par(mfrow=c(1,2))
plot(estR)
plot(estR2)

R_infections <- as.data.frame(estR$R)
R_hospitalizations <- as.data.frame(estR2$R)

# Confidence interval of calculated effective reproduction rates
estR$conf.int
estR2$conf.int
```

```
## REGRESSION MODEL FOR Re  (Section 4.3.3) --------------------------------
## LINEAR MODEL -----------------------------
# Full model normalized
full_linear <- lm(R_calculated~ Rel_hum + Temp_ave + Wind_speed +
                   At_home_average + Traveling_average + Schools + Catering + Events + Facemasks,
                 data = InputDataR)
summary(full_linear) #
vif(full_linear)

# Full model AVERAGED
full_linear <- lm(R_calculated~ R1week + R2week + Temp_ave + At_home_average + Traveling_average +
                   Schools + Events + Facemasks + Catering,
                 data = R2DATA)
summary(full_linear) #
vif(full_linear)

## NON LINEAR MODEL ------------------------------
full_poly <- lm(R_calculated~ Rel_hum + Temp_ave + Wind_speed +
              At_home_average + Traveling_average + Schools + Catering + Events + Facemasks +
             I(Rel_hum^2) + I(Rel_hum^3) + I(Rel_hum^4) + I(Temp_ave^2) + I(Temp_ave^3) + I(Temp_ave^4) +
             I(Wind_speed^2) + I(Wind_speed^3) + I(Wind_speed^4) + I(At_home_average^2) +
             I(At_home_average^3) + I(At_home_average^4) + I(Traveling_average^2)+ I(Traveling_average^3) +
             I(Traveling_average^4) + I(Schools^2) + I(Schools^3) + I(Schools^4) + I(Catering^2)+ I(Catering^3) +
             I(Events^3) + I(Facemasks^3), data = R2DATA)
summary(full_poly) #

backward_poly <- lm(R_calculated ~ Wind_speed + At_home_average + Traveling_average +
              Schools + Events + I(Temp_ave^2) + I(Temp_ave^3) + I(Temp_ave^4) +
              I(Wind_speed^2) + I(Wind_speed^3) + I(Wind_speed^4) + I(At_home_average^2) +
              I(At_home_average^3) + I(At_home_average^4) + I(Traveling_average^2) + I(Traveling_average^3) +
             I(Traveling_average^4) + I(Schools^2) + I(Schools^3) + I(Schools^4) + I(Catering^2) + I(Catering^3) +
             I(Events^3) + I(Facemasks^3), data = R2DATA)
summary(backward_poly) #

## INTERACTION MODEL --------------------------
full_interaction <- lm(R_calculated~ Rel_hum + Temp_ave + Wind_speed +
             At_home_average + Traveling_average + Schools + Catering + Events + Facemasks +
             Rel_hum*Temp_ave + At_home_average*Traveling_average*Events +
             At_home_average*Traveling_average*Schools + At_home_average*Traveling_average*Catering +
             At_home_average*Traveling_average*Facemasks
                       , data = R2DATA) #
summary(full_interaction) #

backward_interaction <- lm(R_calculated ~ Rel_hum + Temp_ave + Wind_speed + At_home_average +
             Traveling_average + Schools + Catering + Events + Facemasks + At_home_average:Traveling_average
             + At_home_average:Events + Traveling_average:Events + Traveling_average:Schools +
             Traveling_average:Catering + At_home_average:Facemasks + Traveling_average:Facemasks +
             At_home_average:Traveling_average:Events + At_home_average:Traveling_average:Facemasks
             , data = R2DATA) #
summary(backward_interaction) #
backward_interaction_significance <- lm(R_calculated ~ Rel_hum  + Wind_speed + At_home_average +
             Traveling_average + Schools + Catering + Events + Facemasks + At_home_average:Traveling_average
             + Traveling_average:Events + Traveling_average:Schools + Traveling_average:Catering +
             At_home_average:Facemasks, data = R2DATA) #
summary(backward_interaction_significance) #


## IDENTIFY NON-LINEARITY; RESIDUAL PLOTS ------------------------
par(mfrow=c(1,2))
plot(predict(full_linear),residuals(full_linear)) #
plot(predict(backward_interaction),residuals(backward_interaction)) #
```

```
# BACKWARD SELECTION ------------------------------
## Backward stepwise selection
step_backward <- stepAIC(full_linear, direction="backward", trace = FALSE)
step_backward$anova
step_backward$coefficients

# BOOTSTRAP BACKWARD ------------------------------------
set.seed(100)
BootBack <- boot.stepAIC(model_R_NORM, data = R2DATA, B = 500, alpha = 0.05,
                         direction = "backward", verbose = T)
BootBack$Significance # shows significance of all variables included in the model
BootBack$OrigStepAIC

## RESIDUAL HISTOGRAMS (Appendix H) -----------------------------
# LINEAR MODEL PLOT
r0histogram <- qplot(full_linear$residuals,
            geom = "histogram",
            bins = 40) +
  labs(title = "Histogram of residuals for linear regression model",
       x = "residual")
r0histogram #

# INTERACTION MODEL PLOT
r0histogramint <- qplot(backward_interaction$residuals,
                     geom = "histogram",
                     bins = 40) +
  labs(title = "Histogram of residuals for interaction regression model",
       x = "residual")
r0histogramint #


## FORMULA CONFIRMED CASES (Section 4.3.4) -----------------------------
# LINEAR TESTING MODEL
Testing <- lm(Confirmed_infected~AverageTests+Besmettelijk_9, data = TestingDATA)
summary(Testing)

# LINEAR TESTING MODEL with LOG response
TestingLOG <- lm(log(Confirmed_infected)~AverageTests+Besmettelijk_9, data = TestingDATA)
summary(TestingLOG)

# INTERACTION TESTING MODEL
TestingInteraction <- lm(Confirmed_infected~AverageTests+Besmettelijk_9+AverageTests*Besmettelijk_9, data =
TestingDATA)
summary(TestingInteraction)

# INTERACTION TESTING MODEL with LOG response
TestingInteractionLOG <-
lm(log(Confirmed_infected)~AverageTests+Besmettelijk_9+AverageTests*Besmettelijk_9, data = TestingDATA)
summary(TestingInteractionLOG)

par(mfrow=c(1,2))
plot(predict(Testing),residuals(Testing)) #
plot(predict(TestingLOG),residuals(TestingLOG)) #
plot(predict(TestingInteraction),residuals(TestingInteraction)) #
plot(predict(TestingInteractionLOG),residuals(TestingInteractionLOG)) #

#  PLOT RESIDUAL HISTOGRAM
r0histogram <- qplot(TestingLOG$residuals,
                     geom = "histogram",
                     bins = 40) +
  labs(title = "Histogram of residuals for testing",
       x = "residual")
r0histogram #
```

```
# MAKE TEST AND TRAINING DATASET
set.seed(1234)
model <- createDataPartition(TestingDATA$Confirmed_infected, p=0.8, list=FALSE)
train <- TestingDATA[ model, ]
test <- TestingDATA[ -model, ]

# TRAIN THE MODEL
training <- lm(Confirmed_infected~AverageTests+Besmettelijk_9, data = train)
summary(training) #

# DETERMINE MSE
testing <- predict(training, newdata = test)
MSE <- mean((testing - test$Confirmed_infected)^2)
MSE #

## VALIDATION (Section 4.3.3)  ---------------------------------------
## MAKE TESTING AND TRAINING DATASET
# Datasets for reproduction rate
set.seed(1234)
modelR <- createDataPartition(R2DATA$R_calculated, p=0.8, list=FALSE)
trainR <- R2DATA[ modelR, ]
testR <- R2DATA[ -modelR, ]

# TRAINED MODELS FOR VALIDATION
# Linear model trained
trainingR <- lm(R_calculated ~ R1week + R2week + Temp_ave + At_home_average + Traveling_average +
            Schools + Events + Facemasks + Catering, data = trainR)
summary(trainingR) #

# Interaction model trained
trainingR <- lm(R_calculated ~ Rel_hum + Temp_ave + Wind_speed + At_home_average +
            Traveling_average + Schools + Catering + Events + Facemasks +
            At_home_average:Traveling_average + At_home_average:Events +
            Traveling_average:Events + Traveling_average:Schools + Traveling_average:Catering +
            At_home_average:Facemasks + Traveling_average:Facemasks +
            At_home_average:Traveling_average:Events + At_home_average:Traveling_average:Facemasks,
data = trainR)
summary(trainingR) #

# Polynomial model trained
trainingR <- lm(R_calculated ~ Wind_speed + At_home_average + Traveling_average +
            Schools + Events + I(Temp_ave^2) + I(Temp_ave^3) + I(Temp_ave^4) +
            I(Wind_speed^2) + I(Wind_speed^3) + I(Wind_speed^4) + I(At_home_average^2) +
            I(At_home_average^3) + I(At_home_average^4) + I(Traveling_average^2) +
            I(Traveling_average^3) + I(Traveling_average^4) + I(Schools^2) +
            I(Schools^3) + I(Schools^4) + I(Catering^2) + I(Catering^3) +
            I(Events^3) + I(Facemasks^3), data = trainR)
summary(trainingR) #

# MSE Reproduction rate
testingR <- predict(trainingR, newdata = testR)
MSER <- mean((testingR - testR$R_calculated)^2)
MSER #
```

```
# BOOTSTRAP OPTIMISM CALCULATIONS ----------------------------------
rsq <- function(x,y) cor(x,y)^2

# OUTCOME MODEL R2
my_boot <- function(data,index){
  bootdata <- data[index,]
  f <- lm(R_calculated ~  R1week + R2week + Temp_ave + At_home_average + Traveling_average +
           Schools + Events + Facemasks + Catering
         , data=bootdata)
  p <- predict(f, newdata = data)
  p_orig <- predict(f,newdata = bootdata)
  MSE <- mean((p - data$R_calculated)^2)
  MSEorig <- mean((p_orig - bootdata$R_calculated)^2)
  Optimism_MSE <- MSE - MSEorig
  r2 <- rsq(p,data$R_calculated)
  r2_orig <- rsq(p_orig,bootdata$R_calculated)
  r2correction <- r2 - r2_orig
  return(c(Optimism_MSE, MSE, MSEorig,r2,r2_orig,r2correction))}

set.seed(2020)
result <- boot(data = R2DATA, my_boot, R=500)
meanOptimism <- mean(result$t[,1])
plot(result, index = 1)
histogram(result$t[,2])  ## histogram AUC bootstrap model
histogram(result$t[,3])  ## histogram AUC bootstrap model op originele data

colnames(result$t) <- c("MSE optimism", "MSE", "MSE orig","r2","r2 orig","r2 correction")
head(result$t)
colMeans(result$t) #

## VALIDATION (Section 4.3.5)  --------------------------------------
# FULL MODELS
# Model for deaths
full_linear <- lm(Deaths_infected~ Temp_ave + At_home_average + Traveling_average +
                   Schools + Events + Facemasks + Catering,
                 data = deathsDATA)
summary(full_linear) #

# Model for hospitalizations
full_linear <- lm(Hospitalized_infected~ Temp_ave + At_home_average + Traveling_average +
                   Schools + Events + Facemasks + Catering,
                 data = hospiDATA)
summary(full_linear) #

## TRAINED MODELS
# DEATHS
# Make training and test dataset
set.seed(1234)
modelD <- createDataPartition(deathsDATA$Deaths_infected, p=0.8, list=FALSE)
trainD <- deathsDATA[ modelD, ]
testD <- deathsDATA[ -modelD, ]

# Linear model trained for deaths
trainingD <- lm(Deaths_infected ~ Temp_ave + At_home_average + Traveling_average +
                 Schools + Events + Facemasks + Catering, data = trainD)
summary(trainingD) #

# MSE for deaths
testingD <- predict(trainingD, newdata = testD)
MSED <- mean((testingD - testD$Deaths_infected)^2)
MSED #
```

```
# HOSPITALIZATIONS
# Make training and test dataset
set.seed(1234)
modelH <- createDataPartition(hospiDATA$Hospitalized_infected, p=0.8, list=FALSE)
trainH <- hospiDATA[ modelH, ]
testH <- hospiDATA[ -modelH, ]

# Linear model hospitalizations trained
trainingH <- lm(Hospitalized_infected ~ Temp_ave + At_home_average + Traveling_average +
            Schools + Events + Facemasks + Catering, data = trainH)
summary(trainingH) #

# MSE hospitalizations
testingH <- predict(trainingH, newdata = testH)
MSEH <- mean((testingH - testH$Hospitalized_infected)^2)
MSEH #

# BOOTSTRAP OPTIMISM CALCULATIONS ----------------------------------
(same approach as previous page)
```