Universidad Rey Juan Carlos

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA DE TELECOMUNICACIÓN

BIOMEDICAL ENGINEERING DEGREE

END OF DEGREE PROJECT

Sensitivity Analysis of a Mathematical Model of COVID-19 Transmission Dynamics

Author: Miguel A. Palacios Candelario
Tutor: Alberto Olivares González

Academic year 2020/2021

Overview

The main objective of this End of Degree Project is the study and analysis of the evolution of the Covid-19 pandemic in the target population of the Community of Madrid, considering a compartmental epidemic model, in which vaccination, social distancing measures, and testing of susceptible individuals are included. More precisely, this compartmental model will be based on the fundamentals of the SIR (Susceptible-Infected-Recovered) model, which consists on a system of non-linear differential equations, and will further be modified in order to extract more complex and valuable information.

In order to perform a more reliable study, a reference model has been developed, which aims to mimic the behavior of the pandemic during the month of March 2021 in the Community of Madrid, in terms of deaths and contagions, for a further purpose of predicting such behavior for the following 3 months, as they are critical for the strategy planning and decision-making towards the summertime, being able to ensure both safety and economic stability.

Therefore, the scope of the project and further conclusions have settled around two methods of experimentation:

- On the one hand, the sensitivity analysis of the afored mentioned mathematical model of the SARS-CoV-2 virus transmission dynamics, including the parametric analysis of pharmacological measures and non-pharmacological measures.
- On the other hand, a stochastic sensitivity analysis, as it is a real-life situation which demonstrates a degree of randomness. Stochastic analysis therefore allows understanding of real-world situations where random behavior is crucial.

All the modeling, calculations and results have been performed and obtained through the *Python* software. More precisely, by using the *odeint* and *lsoda* functions from the *FORTRAN* library *odepack*.

"Mathematics as an expression of the human mind reflects the active will, the contemplative reason, and the desire for aesthetic perfection. Its basic elements are logic and intuition, analysis and construction, generality and individuality."

- Richard Courant.

Contents

List of Figures

List of Tables

Chapter 1

Introduction: *The COVID-19 Pandemic*

1.1 The COVID-19 Pandemic

The 2019 outbreak of COVID-19 in Wuhan (China) rapidly turned into a global pandemic, causing the World Health Organization (WHO) to declare the situation as a public health emergency involving all nations. Up to February 2021, more than one year later, the situation has worsened globally regarding the total infections and deaths, the overwhelming situation of the health system and health professionals, and leading to a devastating economic crisis as never seen before. The arrival of the vaccines and the posterior acceleration in the vaccination rate has implied a noticeable relieve for the healthcare system stress, specially the Intensive Care Units (ICU). Moreover, this pharmacological measure has signified a reduction of the total number of deaths and the infection incidence, being a step closer to the control over the spread of the virus.

According to the definition of the WHO [1], the Coronavirus disease is an infectious disease caused by the recently discovered SARS-CoV-2 virus. The most common effect on infected individuals will be respiratory illness, in most cases recovering without treatment. Nevertheless, vulnerable individuals, including the elderly and those with underlying medical problems like cardiovascular disease, diabetes, chronic respiratory disease, and cancer are more likely to develop serious illness.

Moreover, the high risk of infection has been making it difficult for all nations and governments to control the spread of the virus, emphasizing the importance of basic safety measures, as mention by the WHO, which empathizes the impor-

tance of information as the best way to reduce the transmission of the SARS-CoV-2 virus. This also includes basic hygiene measures like the use of face-masks, cleaning and sanitizing of one's hands, and maintain the social distance recommended by Health Institutions.

All these aspects previously mentioned contribute as a part of the motivation for the development of this End of Degree Project, making it attractive for students, scientists or any professional in the matter, as it is an ongoing real-life scenario which requires for an immediate solution, which will make a huge impact in social health and well-being. Moreover, the understanding of a pandemic and the ways to stop it through different analytic tools will also remark the importance of this type of studies, and will constitute the basis for any future situation.

Ever since the SARS-CoV-2 turned into the number one health-priority for governments, many multidisciplinary professionals had to come together to find a solution to all the problems generated by the pandemic. This involved skilled workers as epidemiologists, physicians, mathematicians, or pharmacologists, among others. Although, the lack of information and experience with modern-days pandemics, along with several bad decisions regarding the management of the circumstances, led to an aggravation of the problem that we are still facing today.

1.2 State of Art

As mentioned before, there is no historical information and studies regarding an outbreak at the level of the COVID-19 pandemic, but similar infectious diseases outbreaks, alongside the numerous amount of COVID-19 studies in the past year have facilitated the documentation for this project. To understand the basis of a pandemic disease and its mathematical foundation, the main tasks of this project have been developed around compartmental models and systems of differential equations, used to simplify the mathematical modeling of infectious disease, where the population is divided into compartments, with the assumption that every individual in the same compartment has the same characteristics.

Thus, on the one hand, the theoretical framework relies on books like *Numerical Methods for Engineers*, by Steven C. Chapra and Raymond P. Canale [2], and publications like *The SIR Model and the Foundations of Public Health*, by H. Weiss [3]. On the other hand, the main studies in which this project has been based on have been developed very recently and almost in parallel in some cases. The first type of studies includes the mathematical model with the same parametrization criterion used in this project, including articles like *Early Transmission Dynamics*

in Wuhan, China, of Novel Coronavirus–Infected Pneumonia, by Qun Li et al. [4], or *Nowcasting and Forecasting the Potential Domestic and International Spread of the 2019-nCoV Outbreak Originating in Wuhan, China: A Modelling Study*, by Joseph T. Wu et al. [5].

Moreover, recent studies like *Uncertainty Quantification of a Mathematical Model of COVID-19 Transmission Dynamics with Mass Vaccination Strategy*, by Alberto Olivares and Ernesto Staffetti [6], incorporates a parametric and stochastic study of the development of the pandemic in Spain, which is considered as the basis for this End of Degree Project.

1.3 Objectives of the Project

In Spain, the most notable way of decision-making during the global pandemic has been relying on the available numerical data on the evolution of the number of deaths and contagions by location. This has determined, for example, the application or deletion of restrictions like the lockdown, the curfew, or the maximum number of people allowed in a closed space. As we can see in Figure 1.1, some of these methods where helpful in the reduction of the total transmissions and deaths.

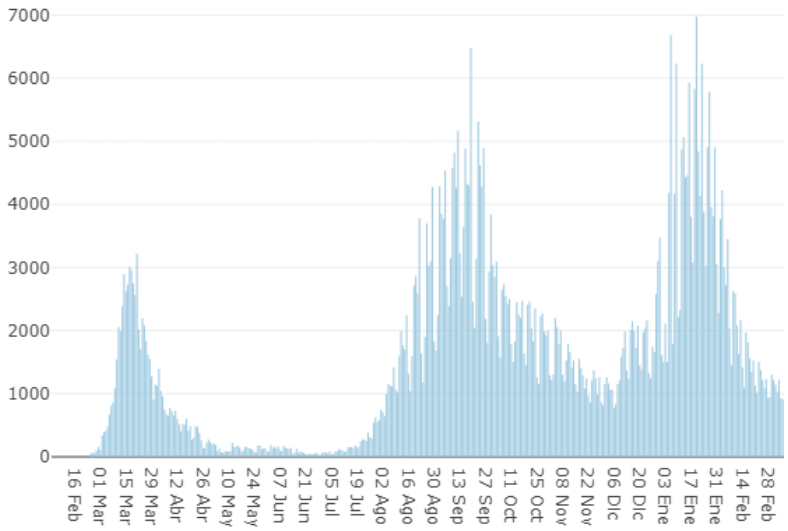


Figure 1.1: Daily new confirmed COVID-19 cases in the Community of Madrid.

The picture has changed dramatically since the beginning the vaccination campaign in January 2021 when there were reportedly over 500 deaths a day. Months

later, at the beginning of May 2021, this number has been reduced to 50. The cumulative incidence rate has began to lower, being most of the contagions among people under 50, who are significantly less likely to become severely ill, or die from COVID-19. This is indeed as a result of vaccination. Both the number of deaths and contagions reflect the increase in the number of people that has been vaccinated in each age bracket.

Nevertheless, there are still many concepts and conclusions that cannot be extracted only from the available data. Furthermore, these pieces of information can be withdrawn from simple mathematical tools and models, which help to lay a theoretical foundation for public health interventions, establishing a basis for decision-making professionals.

In this project, it will be presented the basis of such mathematical models, and how their manipulation and understanding can be useful for working alongside professionals like epidemiologists in the handling of the expansion of the SARS-CoV-2 virus. Moreover, the manipulation of these theoretical formulations will allow us to obtain explanatory results to specific populations.

Therefore, the ultimate goal of the project will be the parametric and stochastic sensitivity analysis of a specific mathematical model developed to represent the COVID-19 transmission dynamics in the Community of Madrid. For this achievement, it is necessary to accomplish the following secondary objectives:

- In the first place, it is necessary to perform a deeper research on previous and parallel works, to obtain the most updated information and data available.
- Understand the functioning of compartmental models, by understanding the basic SIR (Susceptible-Infected-Recovered) model.
- Obtain and understand the results for the SIR model in the Community of Madrid
- Understand a specific improved development of the SIR model, the SEIQR (Susceptible-Exposed-Infected-Quarantined-Recovered) model.
- Understand the effect of the different contingency measures proposed by the governments and health institutions after obtaining the results of the SEIQR model in the Community of Madrid
- Obtain and compare the results from both the SIR and the SEIQR models in the target population of the Community of Madrid.

- Finally, perform a sensitivity analysis, consisting in two parts: analyzing the pharmacological and non-pharmacological measures parametrically, and secondly, a stochastic analysis by introducing Monte Carlo simulations.

1.4 Structure of the Project

This project includes an organized and structured guideline, which covers five different chapters:

- In *Chapter 1*, as seen, the motivation for the project, the state of art, and the main objectives are presented.
- In *Chapter 2: Mathematical Model of COVID-19 Transmission Dynamics*, a brief description is given about the SIR model dynamics and the improved and developed SEIQR model.
- *Chapter 3: Numerical Methods for Ordinary Differential Equations*, includes in the first place, the Runge- Kutta methods for solving ordinary differential equations, like the SIR and the SEIQR model. Posteriorly, this methods are implemented using a *Python* library and adjusted to our code.
- *Chapter 4: Numerical Experiments*, contains everything regarding the numerical experimentation and *Python* simulations, including the parametric sensitivity analysis (divided in pharmacological and non-pharmacological measures), and the stochastic analysis.
- In *Chapter 5: Conclusions*, the final discussion is included, where the main conclusions are presented, and also, the acquired skills during the development of the project, as well as the future lines of investigation possible.
- Lastly, in the annexes and bibliography, there will be the code for the simulations and the sources of information that validate the project, respectively.

Chapter 2

Mathematical Model of COVID-19 Transmission Dynamics

During this chapter, following [3], it will be explained the theoretical basis of the SIR compartmental model, to further describe the more complex version, the SEIQR model introduced by [7] for further computations and analysis of COVID-19 transmission dynamics.

2.1 The SIR Model

The SIR model is a basic transmission deterministic model for a directly transmitted infectious disease, developed by Ronald Ross and William Hamer in the early twentieth century. The model is based on a system of three coupled non-linear ordinary differential equations which does not include an explicit formula solution. Nevertheless, it is still possible to perform calculations in order to extract valuable information.

These types of models are called compartmental models, as they divide the population under study in several compartments or blocks, in this case three blocks for Susceptible, Infected and Removed subjects. The sizes of these blocks are represented as functions of time t by $S(t)$, $I(t)$, and $R(t)$.

The SIR model is based on many strong assumptions, which will be later mathematically demonstrated, regarding topics like the duration of an epidemic, the maximum number of infected individuals, or that an epidemic eventually dies. One of the most notable assumptions made by the model is the mass action mixing, which assumes that the encounter rate between infected and susceptible individuals is directly proportional to the product of their population sizes. Therefore, if the

size of either group is doubled, this would result in twice as many new infections. This assumption implies that the members of both groups are homogeneously distributed. This assumption is reasonable, as most members in a population have a small number of close contacts, like family members, co-workers or classmates.

Figure 2.1: Flow diagram of the SIR model

This well-mixing hypothesis allows the use of Ordinary Differential Equations (ODEs) instead of partial differential equations or agent based models, and therefore being easier to parametrize and analyze. This way, the SIR model can be defined as the following ODE system:

$$\dot{S}(t) = -\beta S(t)I(t), \tag{2.1}$$
$$\dot{I}(t) = \beta S(t)I(t) - \alpha I(t), \tag{2.2}$$
$$\dot{R}(t) = \alpha I(t). \tag{2.3}$$

As seen, the system includes the following two parameters:

- β symbolizes the transmission rate. It is always greater than zero. Moreover, β can be broken down as the product of κ (contacts sufficient for transmission per unit time) and τ (transmissibility of the infectious disease), divided by the total population N, namely:

$$\beta = \frac{\kappa\tau}{N}. \tag{2.4}$$

- α represents the recovery rate of a given infectious disease. It also is always greater than zero. From this information, the duration of the infection D can be deduced:

$$D = \frac{1}{\alpha}. \tag{2.5}$$

Moreover, the system can be simplified into a system of two ODEs, equations (2.1) and (2.2), since

$$R(t) = N - S(t) - I(t). \tag{2.6}$$

From this system of equations, many properties can be proven in order to explain the nature and behavior of a pandemic:

1. **Long term limits exist:** Being the right hand side of equation (2.1) negative and the right hand side of (2.3), and since:

$$0 \leq S(t) \leq S(0) \leq N \quad \text{and} \quad 0 \leq R(0) \leq R(t) \leq N,$$

implies that the following limits hold

$$\begin{aligned} S(\infty) &= \lim_{t \to +\infty} S(t), \\ R(\infty) &= \lim_{t \to +\infty} R(t), \end{aligned}$$

and thus

$$I(\infty) = \lim_{t \to +\infty} I(t) \quad = \quad N - S(\infty) - R(\infty)$$

exist.

2. **The disease always dies out:** It can be observed, as there exist long term limits, that for a sufficiently large t:

$$\dot{R}(t) > \quad \frac{\alpha I}{2} \quad > 0,$$

implying that

$$R(\infty) \quad = \quad \infty,$$

which is a contradiction.

3. **The Epidemic Threshold Theorem:** This theorem holds that there exists a threshold value, called the effective reproductive number, R_e, that will determine if a disease will quickly die out, or on the contrary, spread out and cause a pandemic. On the other hand, the average reproductive number, R_0, represents the number of secondarily infected individuals generated by a primary infected individual. These coefficients can be calculated as

$$R_e \quad = \quad \frac{S(0)}{N} \frac{\kappa \tau}{\alpha}, \tag{2.7}$$

$$R_0 \quad = \quad \frac{\kappa \tau}{\alpha}. \tag{2.8}$$

Thus, combining both equations

$$R_e \quad = \quad \frac{S(0)}{N} R_0 \tag{2.9}$$

(a) If $R_e \leq 1$, then $I(t)$ will decrease to zero as t $\to \infty$. If we take into account that the whole population is susceptible, then

$$\begin{aligned} S(0) &= N-1, \\ I(0) &= 1, \\ R(0) &= 0, \end{aligned}$$

and thus, R_e can be noted as

$$R_e = \frac{N-1}{N}\frac{\kappa\tau}{\alpha},$$

and this way, it can be shown that

$$\dot{I}(t) = \beta S(t)I(t) - \alpha I(t) = (\beta S(t) - \alpha)I(t),$$

where

$$(\beta S(t) - \alpha)I(t) \leq (\beta S(0) - \alpha)I(t),$$

with

$$\begin{aligned} (\beta S(0) - \alpha)I(t) &= \left(\frac{\beta S(0)}{\alpha} - \frac{\alpha}{\alpha}\right) I(t)\alpha = \left[\left(\frac{1}{\alpha}(N-1)\beta\right) - 1\right] I(t)\alpha \\ &= \left[\left(\frac{1}{\alpha}(N-1)\frac{\kappa\tau}{N}\right) - 1\right] I(t)\alpha = \left[\left(\frac{S(0)}{N}\frac{\kappa\tau}{\alpha}\right) - 1\right] I(t)\alpha \\ &= (R_e - 1)I(t)\alpha. \end{aligned}$$

Therefore

$$(R_e - 1)I(t)\alpha \leq 0 \text{ for all } R_e < 1$$

This, along with $I(\infty) = 0$ shown before, proves this point.

(b) If Re > 1, then $I(t)$ will increase until it reaches a maximum and posteriorly it will decrease to zero as t $\to \infty$. This constitutes the well known infected population curve. As seen previously:

$$\dot{I}(0) = (R_e - 1)I(0)\alpha,$$

thus,

$$(R_e - 1)I(0)\alpha > 0 \text{ for all } R_e > 1$$

Therefore, $I(t)$ is increasing at $t = 0$. Moreover, equation (2.2) implies a unique critical point for $I(t)$ and a posterior decrease, as shown in the previous statement.

(c) Lastly, it can follow that an epidemic occurs if $R_0 > 1$, meaning that the transmissibility times the effective contacts, $\kappa\tau$, is greater than the recovery rate, α.

As mentioned previously, the existence of these definitions and parameters will be enough to determine the graphical representation of the famous infection curve. The existence of this threshold of infection, R_e, has been missed out by epidemiologists and other professionals, as it cannot be deducted from data alone. Thus, it requires from mathematical analysis to be extracted and studied.

4. **The Public Health interpretation of the effective reproductive number:** As a deduction from equation (2.7), the effective reproductive number, R_e can be defined as the number of new infections in a population, caused by each infected individual at the beginning of the outbreak. Therefore, it is indeed reinforcing the statement of the first theorem: If the first infected individual infects three susceptible individuals, and each of those people infects another three susceptible individuals, then surely the growth of the infections is going to be exponential.

5. **Public Health interventions:** Looking at equation (2.7), we can now start deducing the theoretical basis for some of the most common governmental measures applied during the outbreak of a pandemic, as we have been able to see recently due to the COVID-19. Therefore, the main goal will be to reduce R_e to less than one:

   (a) By reducing the duration of the disease infection, D, which in this case is complicated as no anti-virals have shown efficacy on this term.

   (b) By reducing the contact rate, κ, from measures as reducing the maximum number of people in a closed space, to more extreme cases as quarantine.

   (c) By reducing the transmissibility, τ, in a way affected by measurements like the ones exposed in the previous point, but also some simpler ways of reducing this factor like wearing face masks, using hand sanitizer, or constant ventilation of a closed area.

   (d) By reducing $S(0)$, which will only be possible through vaccination, which will be explained more in detail in the next point.

6. **The logic behind vaccination and heard immunity:** The main purpose of vaccination, from a SIR point of view, is to transform susceptible individuals into removed ones. The main problem with vaccination is that it is not

always 100% effective, nor the tolerance will be the same for every person (some other diseases or complications may occur). Moreover, mass vaccination carries along a huge expenditure for the healthcare system of a country. Nonetheless, there exists a vaccination threshold or heard immunity, proving that an epidemic can be fought by vaccinating a specific percentage of the population. This value is really powerful, but not so intuitive, as it cannot be discerned from data. We recall that for an epidemic to end, $R_e \leq 1$. We denote as p the amount of susceptible individuals who have been vaccinated, assuming a 100% effectiveness of the drug. Then, the amount of susceptible individuals left can be noted as $(1-p)S(0)$. Therefore:

$$\begin{aligned} R_e &\leq 1, \\ (1-p)\frac{\kappa\tau}{\alpha}S(0) &\leq 1, \\ (1-p) &\leq \frac{1}{\frac{\kappa\tau}{\alpha}S(0)}, \\ 1 - \frac{1}{\frac{\kappa\tau}{\alpha}S(0)} &\leq p. \end{aligned}$$

This way, the vaccination threshold or heard immunity can be bounded as follows:

$$p_c \geq 1 - \frac{1}{R_e}. \tag{2.10}$$

For a vaccine whose effectiveness is not 100%, the vaccination threshold will have to be divided by that specific percentage. For instance, an effective reproductive number, R_e, of 2.2 will result in a vaccination threshold, p_c, of 0.54. Therefore, it would be sufficient to vaccinate 54% of the target population with a vaccine 100% efficient. However, if such efficacy was, for example, 80%, then the heard immunity would be reached by immunizing 68.1% of the population.

7. **The maximum number of infected individuals:** Although it has been mentioned that there is not an explicit solution for the ODE composing the SIR system, there is still solution formula for I_{max}. Dividing equations (2.1) by (2.2) yields the separable differential equation:

$$\frac{dS(t)}{dI(t)} = \frac{-\beta S(t)I(t)}{\beta S(t)I(t) - \alpha I(t)} = \frac{-\beta S(t)}{\beta S(t) - \alpha}.$$

Thus, separating the equations

$$\int \frac{\beta S(t) - \alpha}{\beta S(t)} dS(t) \;=\; -\int dI(t).$$

Solving for the left hand side, we have:

$$\begin{aligned}\int \frac{\beta S(t) - \alpha}{\beta S(t)} dS(t) &= \frac{1}{\beta}\int \frac{-\alpha + \beta S(t)}{S(t)} dS(t) = \frac{1}{\beta}\int \frac{-\alpha}{S(t)} + \beta dS(t) \\ &= \frac{1}{\beta}\left(\int \frac{-\alpha}{S(t)} dS(t) + \int \beta dS(t)\right) = -\frac{\alpha}{\beta}\ln S(t) + S(t) + C.\end{aligned}$$

Therefore,

$$-\frac{\alpha}{\beta}\ln S(t) + S(t) + C \;=\; -I(t).$$

Equally,

$$C \;=\; -I(t) - S(t) + \frac{\alpha}{\beta}\ln S(t).$$

Thus, for every $t \geq 0$,

$$I(t) + S(t) - \frac{\alpha}{\beta}\ln S(t) \;=\; I(0) + S(0) - \frac{\alpha}{\beta}\ln S(0).$$

If we recall from equation (2.2), I_{max} happens when $\dfrac{dI(t)}{dt} = 0$, and consequently, when $\mathrm{S(t)} = \dfrac{\alpha}{\beta}$. Applying it to the previous equation, we can obtain the formula solution for I_{max}:

$$I_{max} + \frac{\alpha}{\beta} - \frac{\alpha}{\beta}\ln\frac{\alpha}{\beta} \;=\; I(0) + S(0) - \frac{\alpha}{\beta}\ln S(0),$$

$$I_{max} \;=\; I(0) + S(0) - \frac{\alpha}{\beta}\ln S(0) - \frac{\alpha}{\beta} + \frac{\alpha}{\beta}\ln\frac{\alpha}{\beta}. \tag{2.11}$$

8. **Why do epidemics end?:** As a continuation to point 2 (The disease always dies out), there has always been an uncertainty regarding why epidemics end. For a long time, it has been thought that the main reason behind this end is

that there are no longer susceptible individuals to be infected. This can be proven to be false. Dividing equation (2.1) by equation (2.3):

$$\frac{dS(t)}{dR(t)} = \frac{-\beta S(t) I(t)}{\alpha I(t)} = \frac{-\beta S(t)}{\alpha}.$$

Thus, separating the equations

$$\int \frac{1}{S(t)} dS(t) \;=\; -\int \frac{-\beta}{\alpha} dR(t).$$

Integrating for both sides, we obtain

$$\begin{aligned} \ln\left(S(t) - S(0)\right) &= -\frac{\beta}{\alpha}(R(t) - R(0)), \\ \frac{\ln S(t)}{\ln S(0)} &= -\frac{\beta}{\alpha}(R(t) - R(0)). \end{aligned}$$

eading to

$$\ln S(t) \;=\; -\frac{\beta}{\alpha}(R(t) - R(0)) \ln S(0).$$

Applying logarithmic properties on both sides of the equation:

$$S(t) \;=\; S(0) \exp\left(-\frac{\beta}{\alpha}(R(t) - R(0))\right).$$

Since $0 \leq R(t) - R(0) \leq N$, then $S(t) \geq S(0) \exp\left(-\frac{\beta}{\alpha}(R(t) - R(0))\right)$, and therefore:

$$S(\infty) \geq S(0) \exp\left(-\frac{\beta}{\alpha}(R(t) - R(0))\right) \;=\; S(0) \exp\left(-R_0\right) > 0$$

As seen, an epidemic will end due to the lack of infected individuals, rather than susceptible ones. This statement cannot be deduced from data, being therefore in need of mathematics to be expressed.

These are the basic points for the understanding of a simple SIR model. In practice, it can be parametrized with real data in order to obtain numerical conclusions, although this process is not always easy. In Figure 2.2, it can be observed the basic representation for the solution of the SIR model. According to the calculations, the maximum peak of the infected curve would be reached in day 48, affecting 869235 individuals. A particular scenario has been chosen for this computation:

- The object population chosen is the Community of Madrid at March 1_{st} 2021 and graphed for the following 90 days, counting with a population of 6,7 million people [8]. Therefore, it has been set that $S(0) = 6779888$.
- The initial number of infected individuals, $I(0) = 8143$, have been chosen according to the government's publication at this date [9], as well as the total deaths and total cases for the calculation of $R(0) = 578770$.
- The effective contact rate, $\beta = 0.3$, has been used according to estimations presented in [10], as well as the recovery rate, $\alpha = 1/7$, according to [4].

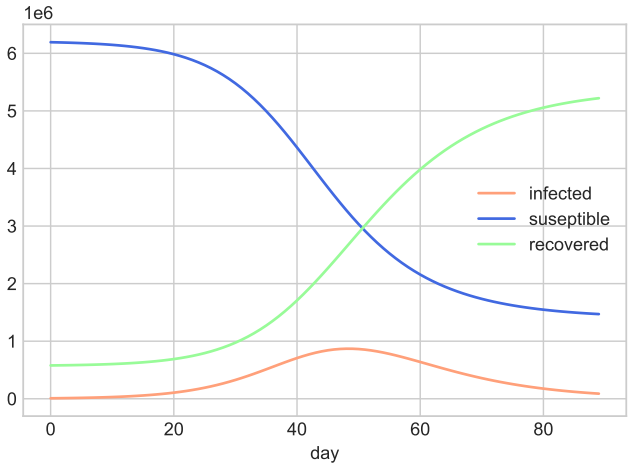


Figure 2.2: SIR model curves for the Community of Madrid assuming $R_0 = 2.1$.

As seen, this mathematical model can bring along a lot of information, nondeductible from data, and that can serve as a basis for the study of a disease expansion, contingency plans or safety measures, among others. Now, a more complex version of this model, the SEIQR model, is going to be introduced, before the posterior analysis of its sensitivity to the values of the parameters.

2.2 The SEIQR Model

The SEIQR model, a modified version of the SIR model, is a nonlinear ODE system, which is used to represent more realistic transmission dynamics [18]. This model includes two new blocks, Exposed and Quarantined, to portray the latent or incubation period of the virus, as well as the possibility of isolate the infected

individuals through quarantine. So far, this model is generic to any viral disease.

For the specific case of the COVID-19 SEIQR model, one more specification has to be made. As seen recently, the virus can affect the population in several ways, from presenting no symptoms at all to mild or severe symptoms, causing hospitalization in some cases. Therefore, there lies the final division of this model in 6 compartments: susceptible, $S(t)$, exposed, $E(t)$, symptomatic infectious, $I_s(t)$, asymptomatic infectious, $I_a(t)$, quarantined, $Q(t)$, and removed, $R(t)$, individuals.

One of the advantages of computing compartimental systems of equations from the SIR model, is that more blocks can be paramtetrized and added to the ODEs, as it is the case of vaccinations or vaccinated individuals more precisely, which would go from the Susceptible block to the Removed one. Although, computations regarding vaccinations will be included posteriorly. Therefore, the sensitivity of the SEIsIaQR model to the parameters will be discussed and computed for two hypothetical situations: whether vaccination happens or not.

In Figure 2.3, we can observe the relationship between the compartmental blocks of the SEIsIaQR model defining a target population, including the action of vaccination.

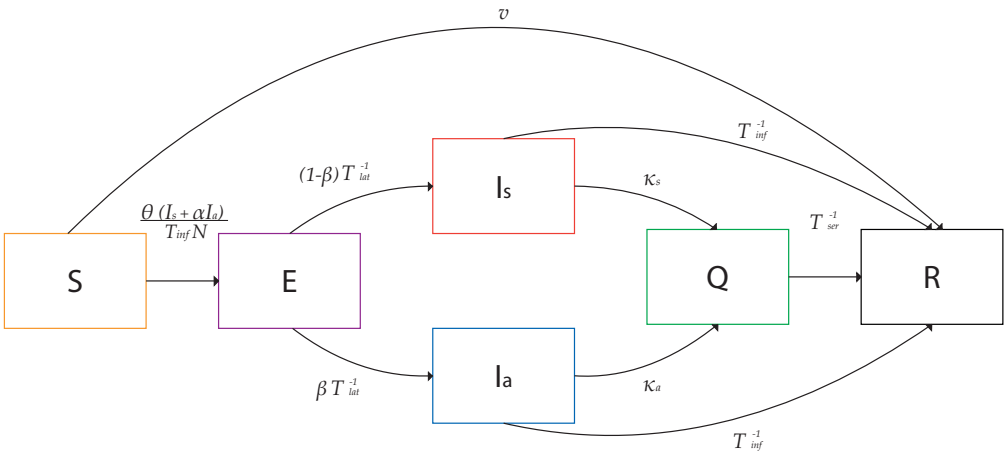


Figure 2.3: Flow diagram of the SEIsIaQR model with vaccination term.

2.2.1 The SEIsIaQR Model Without Vaccination

Based on Figure 2.3, and skipping the vaccination factor for now, the corresponding ODE system can be described as:

$$\dot{S}(t) = -\frac{\theta}{T_{inf}}\frac{I_s(t)+\alpha I_a(t)}{N(t)}S(t), \tag{2.12}$$

$$\dot{E}(t) = \frac{\theta}{T_{inf}}\frac{I_s(t)+\alpha I_a(t)}{N(t)}S(t) - \frac{E(t)}{T_{lat}}, \tag{2.13}$$

$$\dot{I}_s(t) = (1-\beta)\frac{E(t)}{T_{lat}} - \left(\kappa_s + \frac{1}{T_{inf}}\right) I_s(t), \tag{2.14}$$

$$\dot{I}_a(t) = \beta\frac{E(t)}{T_{lat}} - \left(\kappa_a + \frac{1}{T_{inf}}\right) I_a(t), \tag{2.15}$$

$$\dot{Q}(t) = \kappa_s I_s(t) + \kappa_a I_a(t) - \frac{Q(t)}{T_{ser}}, \tag{2.16}$$

$$\dot{R}(t) = \frac{I_s(t)+I_a(t)}{T_{inf}} + \frac{Q(t)}{T_{ser}}. \tag{2.17}$$

Moreover, the total number of confirmed cases can be computed as:

$$\dot{C}(t) = (1-\beta)\frac{E(t)}{T_{lat}} + \kappa_a I_a(t). \tag{2.18}$$

While the SIR system counted with 3 variables and 2 parameters, the SEIsIaQR model is more complex, incorporating 3 more variables and having a total of 8 parameters:

1. Variable $S(t)$ represents the number of individuals susceptible to infection, as in the SIR model.

2. Variable $E(t)$ is the number of exposed individuals. This is, those individuals who have been infected, but do not show symptoms yet. At this stage, the disease can be infectious or non infectious.

3. Variable $I_s(t)$ represents the number of symptomatic individuals. They are indeed confirmed cases, and assumed to be showing symptoms.

4. Variable $I_a(t)$ represents the number of asymptomatic individuals, including the mild symptomatic cases.

5. Variable $Q(t)$ are those individuals who have been quarantined due to safety measures.

6. Lastly, variable $R(t)$ are those individuals who have either diseased or overcame the disease. These last ones are considered immune. Vaccinated individuals are also included in this block if a vaccination term is added to the model.

Just as deduced in the SIR model, the system can be simplified, obtaining variable $N(t)$, which represents the total population, satisfying the condition

$$N(t) = S(t) + E(t) + I_s(t) + I_a(t) + Q(t) + R(t), \quad \text{for all } t \in [0, t_F],$$

where t_F is the final time of the time period considered within the system of equations. Now, regarding the parameters, it is important to note that it is not the goal of this project to estimate these parameters, but to modify them within a realistic range of values to analyze their impact on the system of equations, and therefore, their sensibility. We find two type of parameters:

1. Parameters that are independent from mitigation measures. In the case of the SEIsIaRQ model as seen in Figure 2.3, we find:

   (a) T_{ser}, which is a serial interval time, representing the mean time between consecutive transmissions of COVID-19. Following [4], this value has been set to $T_{ser} = 7.5$ days.

   (b) T_{lat} stands for latency time period, or mean incubation period, which is the time interval for which exposed individuals are infectious. This value has been considered to be $T_{lat} = 5.2$ days [5].

   (c) T_{inf} will be the time interval in which an individual really is infectious, calculated as $T_{inf} = T_{ser} - T_{lat}$.

   (d) α represents the ratio of infectiousness between infectious asymptomatic individuals, I_a, and infectious symptomatic, I_s. The value for α has been set to 1, following [7].

   (e) Lastly, β is the population ratio that remains asymptomatic, with a value of $\beta = 0.8$, following [11].

2. The second type of parameters represent the non-pharmacological mitigation actions. These values will be modified to study the sensitivity of the model. These non-pharmacological interventions vary between countries, and even between cities, but they include social distancing (which restrict large gatherings, specially in inner closed spaces), border closures, restaurants and bars closures, and even districts to cities lock-downs of populations, banning any sort of inner traveling except for essential cases:

(a) θ will be the replication factor, being $\theta = \gamma R_0$, where γ is the interaction factor between individuals, and R_0 is the previously described basic reproduction number.

(b) κ_s represents the isolation rate of symptomatic individuals. These value will be closer to one, as the detection of such individuals is not so complicated.

(c) On the other hand, κ_a will be the isolation rate for asymptomatic individuals. Conversely to κ_s, κ_a will have a value closer to 0, as it is harder to identify those individuals who do not present symptoms.

2.2.2 The SEIsIaQR Model With Vaccination

Again, and as seen on Figure 2.3, the equations for the SEIsIaQR model with vaccination term will include just one additional parameter:

$$\dot{S}(t) = -\frac{\theta}{T_{inf}} \frac{I_s(t) + \alpha I_a(t)}{N(t)} S(t) - vS(t), \tag{2.19}$$

$$\dot{E}(t) = \frac{\theta}{T_{inf}} \frac{I_s(t) + \alpha I_a(t)}{N(t)} S(t) - \frac{E(t)}{T_{lat}}, \tag{2.20}$$

$$\dot{I}_s(t) = (1-\beta)\frac{E(t)}{T_{lat}} - \left(\kappa_s + \frac{1}{T_{inf}}\right) I_s(t), \tag{2.21}$$

$$\dot{I}_a(t) = \beta\frac{E(t)}{T_{lat}} - \left(\kappa_a + \frac{1}{T_{inf}}\right) I_a(t), \tag{2.22}$$

$$\dot{Q}(t) = \kappa_s I_s(t) + \kappa_a I_a(t) - \frac{Q(t)}{T_{ser}}, \tag{2.23}$$

$$\dot{R}(t) = \frac{I_s(t) + I_a(t)}{T_{inf}} + \frac{Q(t)}{T_{ser}} + vS(t). \tag{2.24}$$

Therefore, the only significant changes are given on equation (2.19) and equation (2.24), where the proportion of susceptible individuals immunized through vaccination, $vS(t)$, are incorporated into the recovered block.

Previously, two types of parameters were described. The vaccination factor, v, would constitute a third kind of parameter, a pharmacological mitigation action parameter. Therefore, v represents the vaccination rate, which is assumed to be constant, at a value of 0.0015, according to the Spanish Government's vaccination plan and strategy [12]. Moreover, the vaccine efficacy has been considered to be 95%, as both, the *Moderna* and *Pfizer* vaccines have shown this efficacy [13].

For the implementation of the numerical experiments using the SEIsIaQR model

and their posterior analysis, a Python code has been developed, including the ODE system described in both Sections 2.2.1 and 2.2.2, the values for the different parameters, and the time range for which the study is carried on, which applies in this case for the months of March, April and May 2021.

The outcome of this code will include a data-frame plus the values for each block (the amount of people included) from the first to the last day of the experimentation, which allows the curves for the six blocks to be represented.

Chapter 3

Numerical Methods for Ordinary Differential Equations

In this chapter, the solving techniques for the ODEs presented in the previous chapter are shown. This includes: the theoretical basis of the methods for solving ODEs and the Python code developed to obtain the results later discussed.

For a better understanding of how works the *Python* function posteriorly used, it is important to establish a theoretical basis regarding numerical methods for solving ODEs. These numerical methods are the set of techniques used to find numerical approximations to the solutions of ODEs. There is a large selection of methods for the study of ODEs, partially due to the complexity and possibilities of the matter, since even the simplest non-linear models are very difficult to solve with analytic formulas. Therefore, these methods provide with approximate results of the desired solution.

In this case, we are going to focus on the non-stiff methods [2]. Alongside the known single-step methods, they are the backbone of the Runge-Kutta methods for solving ODEs, included in the *Python* modules.

The non-stiff methods are the simplest family of techniques for solving differential equations of the type

$$\frac{dy}{dx} = f(x, y), \tag{3.1}$$

where x is the independent variable and y the dependent variable. In mathematical terms, the numerical approximation of this equation can be expressed as

$$y_{i+1} = y_i + \phi h, \tag{3.2}$$

where ϕ is the estimated slope, used to extrapolate from a previous value y_i to a new one y_{i+1}, over a distance h. This formula is applied step by step, to obtain the analytical trajectory of the solution.

There are two main non-stiff methods: the Euler's method and the Runge-Kutta method. They are both based on the general formula (3.2), but they differ in the way of estimation of the slope ϕ.

3.1 The Euler's Method

In this method, the first derivative yields an estimation of the slope ϕ in x_i such that

$$\phi = f(x_i, y_i), \tag{3.3}$$

where $f(x_i, y_i)$ is the differential equation evaluated in x_i and y_i. Therefore, applying this to the general formula (3.2):

$$y_{i+1} = y_i + f(x_i, y_i)h. \tag{3.4}$$

As seen, the Euler's method establishes a simple way of predicting a point y by using the value of the slope, and extrapolates it over the distance h.

One of the main drawbacks of the Euler's method is its errors implication. Firstly, truncation errors appear due to the approach used to approximate the value of y. This type of error can, at the same time, produce an error in a specific point in time, called local error, or propagate such error to the posterior steps, producing a propagated error. Moreover, deeper insight of truncation errors can be gained by deriving the Eulers's equation directly from the expansion of the Taylos series.

If the function describing the behavior of y obtained as a solution has continuous derivatives, it can be represented by a Taylor series expansion at a starting value:

$$y_{i+1} = y_i + y_i'h + \frac{y_i''}{2!}h^2 + ... + \frac{y_i^{(n)}}{n!}h^n + R_n, \tag{3.5}$$

where

$$h = x_{i+1} - x_i, \tag{3.6}$$

and R_n can be defined as the remainder term, noted as

$$R_n = \frac{y^{(n+1)}\xi}{(n+1)!}h^{n+1}, \tag{3.7}$$

where ξ lies at some point in the interval from x_i to x_{i+1}. Another alternative to reach this solution comes from inserting equation (3.4) into equations (3.5) and (3.6), which yields

$$y_{i+1} = y_i + f(x_i, y_i)h + \frac{f'(x_i, y_i)}{2!}h^2 + ... + \frac{f^{(n-1)}(x_i, y_i)}{n!}h^n + O(h^{n+1}), \quad (3.8)$$

where $O(h^{n+1})$ indicates that the local truncation error is directly proportional to the step size raised to the $(n+1)th$ power.

Therefore, if equations (3.4) and (3.7) are compared, it can be seen that the Euler's method corresponds to the Taylor series, including the term $f(x_i, y_i)h$. Moreover, this comparison indicates that using a finite number of terms from the Taylor series to approximate the solution leads to a truncation error. It is therefore truncated, or left out, a part of the true solution. Furthermore, from subtracting equations (3.4) and (3.7), it is obtained the formula for the local truncation error:

$$E_t = \frac{f'(x_i, y_i)}{2!}h^2 + ... + O(h^{n+1}). \quad (3.9)$$

One of the limitations found in this method is that the Taylor series only provides an estimate of the local truncation error, that is, the error created during a single step of the method. It will not give a measure neither of the propagated error and, as a result, the global truncation error. Nonetheless, the Taylor series expansion still provides useful information about the Euler's method. It can be seen that the local error is proportional to the square of the step size and the first derivative of the differential equation. It can also be shown that the global truncation error is $O(h)$, that is, it is proportional to the step size. Therefore, from these statements, it can be concluded that reducing the step size will also reduce the error. Moreover, the method will give a prediction without errors if the solution of the differential equation is linear, since for a straight line the second derivative will be zero. Therefore, Euler's method is commonly referred to as a first-order method.

One common way to reduce the error of Euler's method can be performed by including higher-order terms of the Taylor series expansion in the solution. For example, including the second-order term from equation (3.8) yields

$$y_{i+1} = y_i + f(x_i, y_i)h + \frac{f'(x_i, y_i)}{2!}h^2, \quad (3.10)$$

where the local truncation error is

$$E_a = \frac{f''(x_i, y_i)}{6}h^3. \quad (3.11)$$

Even though incorporating higher-order terms is sufficiently simple to implement for polynomials, their addition is not so trivial when the ordinary differential equation is more complex. Specifically, ordinary differential equations that are a function of both the dependent and independent variable require the application of the chain rule for its differentiation. As a consequence, alternative one-step methods have emerged. These only require the calculation of first derivatives.

The first of these improvements is the Heun's method. This technique is based on the refinement of the estimate of the slope by determining two derivatives for the interval. Recalling the Euler's method, the slope at the beginning of the interval can be expressed as

$$y_i' = f(x_i, y_i) \tag{3.12}$$

This can be linearly extrapolated to y_{i+1}:

$$y_{i+1}^{(0)} = y_i + f(x_i, y_i)h \tag{3.13}$$

This point would be the stopage point for the ordinary Euler´s method. However, in Heun's method, this point is an intermediate prediction, rather than a final answer. This is why this term is distinguished with a superscript 0. Equation (3.13) is called a predictor equation, and it provides an estimate of y_{i+1} that allows the computation of an approximated slope at the end of the interval:

$$y_{i+1}' = f(x_{i+1}, y_{i+1}^{(0)}). \tag{3.14}$$

Thus, an average slope can be obtained for the interval by combining equation (3.12) and equation (3.14):

$$\dot{y}' = \frac{y_i' + y_{i+1}'}{2} = \frac{f(x_i, y_i) + f(x_{i+1}, y_{i+1}^{(0)})}{2}. \tag{3.15}$$

Therefore, by using this slope to linearly extrapolate from y_i to y_{i+l} using Euler's method, the so called corrector equation is obtained:

$$y_{i+1} = y_i + \frac{f(x_i, y_i) + f(x_{i+1}, y_{i+1}^{(0)})}{2}h \tag{3.16}$$

Thus, the Heun's method is said to be a predictor-corrector approach.

The second improvement to the Euler's method is the Midpoint method. This technique adopts Euler's method to predict a value of y at the midpoint of the interval:

$$y_{i+\frac{1}{2}} = y_i + f(x_i, y_i)\frac{h}{2}. \tag{3.17}$$

Then, the slope at the midpoint can be calculated with this predicted value, such that:

$$y'_{i+\frac{1}{2}} = f(x_{i+\frac{1}{2}}, y_{i+\frac{1}{2}}). \tag{3.18}$$

This is presumed to constitute an accurate approximation of the average slope for the entire interval. This slope is then used to extrapolate linearly from x_i to x_{i+l}:

$$y_{i+1} = y_i + f(x_{i+\frac{1}{2}}, y_{i+\frac{1}{2}})h. \tag{3.19}$$

The Midpoint method is, compared to the Euler's method, more reliable, as it incorporates an estimate of a slope at the midpoint of the interval to predict. Recalling equation (3.18), this equation contains a local truncation error of $O(h^2)$, comparing with the forward approximation developed in Euler's method, which has an error of $O(h)$, as seen previously. Hence, the local error of the Midpoint method is $O(h^3)$, and the global error is $O(h^2)$.

3.2 The Runge-Kutta Methods

The Runge-Kutta methods aim to reach the Taylor series approach accuracy without the need of calculating higher derivatives. There exist several variations of the Runge-kutta methods, but they all lie in the general form of:

$$y_{i+1} = y_i + \phi(x_i, y_i, h)h, \tag{3.20}$$

where $\phi(x_i, y_i, h)$ is known as the increment function, which can be understood as a characteristic slope over the interval. The increment function can be written in general form as

$$\phi = a_1k_1 + a_2k_2 + ... + a_nk_n, \tag{3.21}$$

where the $a_i, i = 1, \ldots, n$, are constants, and the $k_n, i = 1, \ldots, n$, are

$$k_1 = f(x_i, y_i), \tag{3.22}$$
$$k_2 = f(x_i, p_1h, y_i + q_{11}k_1h), \tag{3.23}$$
$$k_n = f(x_i, p_{n-1}h, y_i + q_{n-1,1}k_1h + ... + q_{n-1,n-1}k_{n-1}h), \tag{3.24}$$

being p and q constants. It has to be noted that each equation appears in the next one. This way, k_1 appears in the equation for k_2 and so on. Therefore, since each k_i is a functional evaluation, the Runge-Kutta methods are very efficient for computer calculations.

Several types of Runge-Kutta methods can be obtained by applying different numbers of terms in the increment function as specified by n. Thus, first-order Runge-Kutta method with $n = 1$ is the Euler's method. Once n is chosen, values for the a, p, and q are assessed by equaling equation (3.20) to the corresponding terms in a Taylor series expansion. Hence, the number of terms, n, normally represents the order of the approach.

In the case of the *Python* code, using the *lsoda* function from the library *odeint*, the algorithm applied is of adaptive type, which means that it adapts to the step size. It estimates the integration error and this estimate is then used to adapt the step size such that the estimated error is below a specified threshold. In particular, the order of the Runge-Kutta method included in the *odeint* function is by default 3.

The result of the derivation applied with an order 3 Runge-Kutta method is six equations with eight unknowns. Therefore, values for two of the unknowns must be specified a priori in order to determine the remaining parameters. In this case, the error term $O(h^4)$ will solve up to the third order term in the Taylor series expansion, being in this case

$$y_{i+1} = y_i + \frac{1}{6}(k_1 + 4k_2 + k_3)h, \tag{3.25}$$

where

$$k_1 = f(x_i, y_i), \tag{3.26}$$

$$k_2 = f(x_i + \frac{1}{2}h, y_i + \frac{1}{2}h), \tag{3.27}$$

$$k_3 = f(x_i + h, y_i - k_1 h + 2k_2 h). \tag{3.28}$$

Third-order Runge-Kutta methods have local error $O(h^4)$, and the global error is $O(h^3)$. Moreover, this yields to an exact results when the solution is a cubic. When there are polynomial functions involved, equation (3.21) to equation (3.28) will also be exact when the differential equation is cubic and the solution is quartic.

3.3 The ODEs Solver in Python

The basis of the *Python* module, which numerically solves ODEs, relies on the function *odeint*, which solves the initial value problem for stiff or non-stiff systems of ordinary differential equations, using the *lsoda* function from the *FORTRAN* library *odepack* [14]. This function, written by the mathematicians Linda R. Petzold and Alan C. Hindmarsh, solves systems of the type (3.1), with a dense or banded Jacobian matrix when the problem is stiff, but it automatically selects between non-stiff and stiff methods, choosing the most adequate. It uses the non-stiff method initially, and dynamically monitors data in order to decide which method to use. This means that the user does not have to determine whether the problem is stiff or not, and the solver will automatically choose the appropriate method.

Chapter 4

Numerical Experiments

4.1 First Approach: Reference Model

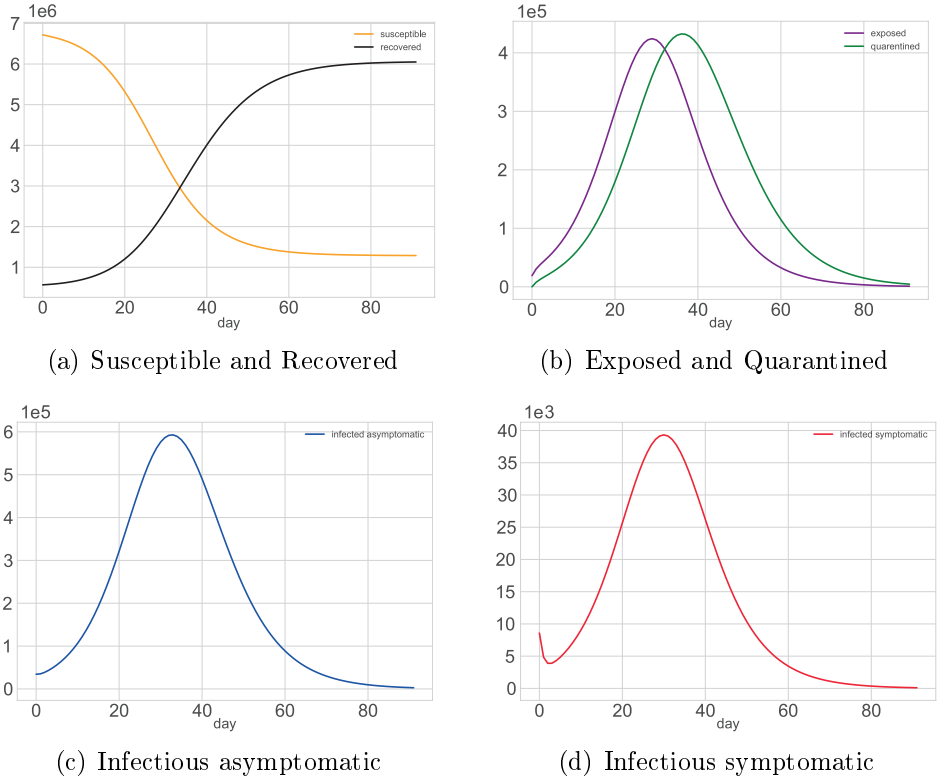
During the computation of the first scenarios for the sensitivity analysis of the model, a *Python* code including the previously explained *Odeint* function has been used (see Annex A for more details). The SEIsIaQR reference model based on the evolution of the pandemic during the month of March 2021 has been formulated and solved, when 40829 contagions were reported [9], [15]. This way, comparisons and analysis will be both, easier and more reliable. For the representation of the model, a time period of 92 days has been considered, being day $t = 0$ the 1st of March, and day $t = 91$ will be then the 31st of May. This time period has been chosen for several reasons: Firstly, and as mentioned before, by the time of computing the model scenarios, the data for the month of March has been published, allowing a more realistic comparison and argumentation of the contingency measures proposed later in this document. Secondly, it is an important trimester for evaluating the situation prior to the summer months, as this season is critical for trying to minimize the economic impact of the COVID-19 in Spain, as well as the sanitary impact, as these summer months are historically characterized by vacation displacements, more social interaction, and big gatherings, among others. Thus, proper measures before this season are essential to reduce the number of contagions.

For the development of this reference model, the values of T_{ser}, T_{lat}, T_{inf}, α, and β of the SEIsIaQR model are considered as fixed values, as specified previously in section 2.2. For the non-pharmacological action parameters, the following values have been set:

- γ has been set to 2. This interaction factor represents isolation and social interactions. Therefore, larger values of θ will be equivalent to less social

distancing measures.

- κ_s has been set to 0.8. A complete detection of the symptomatic cases would imply a value of 1. Although, we know this is not true, as the different tests for detection have different sensitivities and specificities, and the false positive rate can go up to 20%, depending on the day when the test is taken and when the symptoms appear [16].
- Lastly, κ_a will have a value of 0.1. This means that, hypothetically, 10% of the asymptomatic cases are detected. This is a difficult task, as this value relies mostly on massive testing.



(a) Susceptible and Recovered

(b) Exposed and Quarantined

(c) Infectious asymptomatic

(d) Infectious symptomatic

Figure 4.1: Number of individuals in each compartment of the SEIsIaQR reference model.

Regarding the population set up:

- As for Figure 2.2, the total population is $N = 6779888$.
- The number of initial infectious symptomatic individuals $I_s(0)$ has been chosen following the government's publication up to date 1st of March, being

$I_s(0) = 8578$ for the diagnosed cases presenting syptoms for the last 14 days [9].

- The exposed individuals variable will have an initial value of $E(0) = 18963$, following the government's publication for total diagnosed cases in the previous 14 days [9].
- Lastly, $Q(0)$ and $R(0)$ are initially 0 for easier visualization purposes.

In Figure 4.1 we can observe the SEIsIaQR reference model representation. As seen in Figure 4.1(c) and Figure 4.1(b), the amount of infected asymptomatic individual is greater than the amount of quarantined individuals up until day 46, and the infected symptomatic individuals reach a peak of 39319 individuals at day 30, as shown in Figure 4.1(d).
This does not differ much from the actual values regarding those dates, as 40829 infections occurred during the month of March 2021 [9], [15]. Moreover, in Figure 4.2 it can be observed the prevalence counts during the 3 month period, reaching a number of almost 3 million contagions during this time.

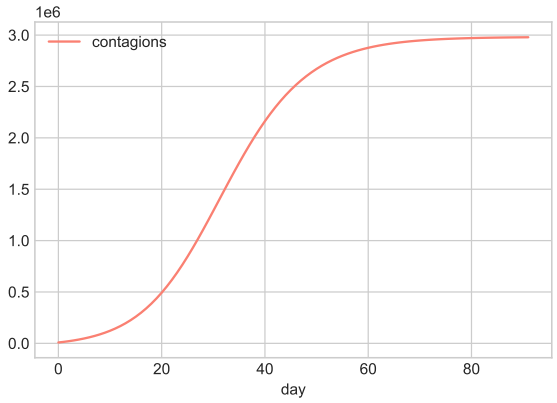


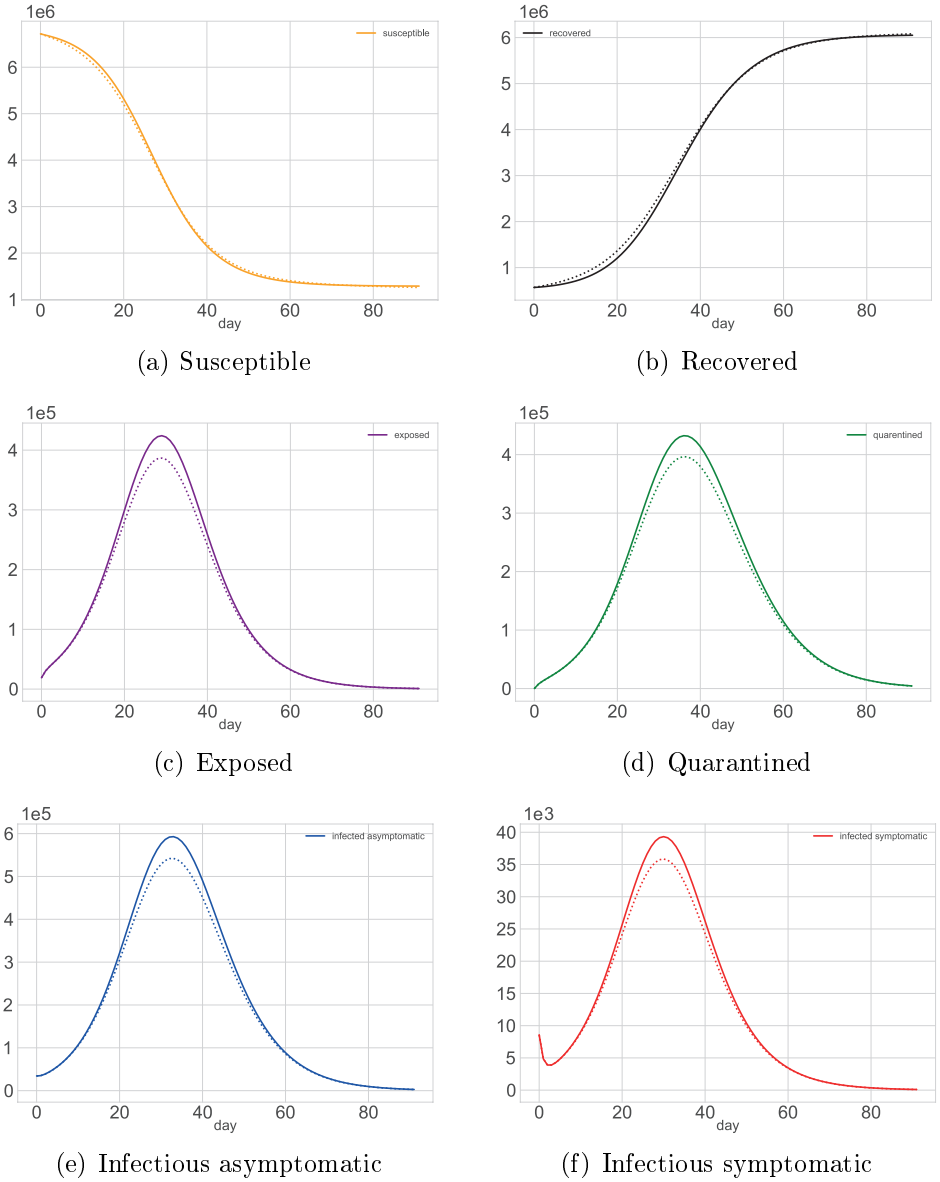
Figure 4.2: Cumulative number of contagions in the SEIsIaQR reference model.

4.1.1 Vaccination Model Vs. Reference Model

At the time of writing this project, mass vaccination is a reality and progressively increasing in pace in the Europen Union, as more vaccines are being distributed. Therefore, the usage of a model which includes this vaccination factor, as shown in Figure 2.3, and described in Section 2.2.2, would make a more complete description and analysis of the problem. For this matter, the Reference Model has been

modified, including the vaccination factor $v = 0.0015$, and has been solved using the same initial values.

In Figure 4.3 it can be seen in detail the comparison between the Vaccination Model (dotted line) and the previously introduced Reference Model (continuous line).



(a) Susceptible (b) Recovered

(c) Exposed (d) Quarantined

(e) Infectious asymptomatic (f) Infectious symptomatic

Figure 4.3: Vaccination Model vs. Reference Model: number of individuals in each compartment.

The administration of the vaccine during the month of March 2021 does not bring a noticeable impact, mainly due to the vaccination rate (0.0015), which is not sufficient to stop the pandemic. This can be observed mainly in Figure 4.3(a) and 4.3(b), where the difference between the susceptible and recovered individuals from both models is barely noticeable. Although, there exists slightly more contrast in the following figures. The number of predicted infectious asymptomatic individuals shown in Figure 4.3(e) is reduced in comparison to the Reference Model, but it is still excessively greater in its maximum (542130 infected asymptomatic individuals) than the quarantined individuals (Figure 4.3(b)) at that day. Figure 4.3(f) shows the infectious symptomatic individuals curve. At its maximum, there is a reduction of an 8.82%, being the maximum at 35853 individuals.
Therefore, the presence of vaccination can suppose a significant change in the evolution of a pandemic if its administration and efficiency reach high enough values, as it can be observed in Figure 4.4, where the total number of contagions is significantly reduced. This topic will be explored in detail later in this section.

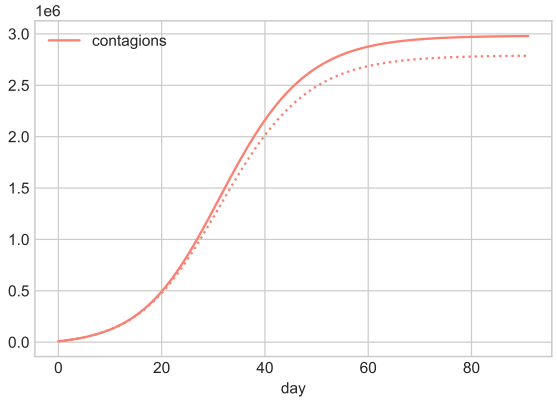


Figure 4.4: Vaccination Model vs. Reference Model: cumulative number of contagions.

From now on, all the experimentation carried along in the next sections will be supported and compared to the Vaccination Model, as it represents a more realistic scenario, since mass vaccination is already established and currently increasing in the target population of the Community of Madrid, with the emergence of new vaccines and the higher supply provided. Thus, the Vaccination Model will be referred as Vaccination Reference Model. Moreover, the figures for the total contagions will be omitted, as the changes in the number of symptomatic and

asymptomatic infected individuals will provide with valuable enough information regarding the level of expansion of the virus, thus plotting contagions will not be necessary.

As for the numerical results of this first approach of the experimentation, the main statistics of the outcome of such simulation have been displayed in Table 4.1 from the 1st of March to the 31st of May. Note that the data for the susceptible and recovered individuals are not shown, as it does not provide with much meaningful information for the matter at this level.

	E	I_s	I_a	Q	**Day**
mean	127429.014407	12015.050941	191490.294425	149000.502218	45.50000
std	131408.425705	12103.111643	184344.623140	135307.037624	26.70206
min	918.998717	97.844912	2711.778853	0.000000	0.00000
25%	12661.808147	1342.672913	33914.888813	27227.109099	22.75000
50%	72016.903345	7116.506295	117555.824509	100359.367869	45.50000
75%	226979.085555	21225.726170	340235.386656	265628.859104	68.25000
max	386642.626503	35853.278758	542130.262414	396568.752708	91.00000

Table 4.1: Main statistics of the Vaccination Reference Model.

As seen, this first simulation is satisfactorily accurate, as 40829 contagions occurred during the month of March according to the Government's publication [9], [15], whereas the maximum number of infected symptomatic individuals from the mentioned simulation is 35853 as seen in Table 4.1. Moreover, the maximum number of affected individuals for each block can be seen in the *max* row in Table 4.1. This information, in combination with the curves presented previously, is very useful to establish a numerical approximation of the events.

4.2 Sensitivity Analysis: Parametric Approach

The following section will be divided into two parts: In the first place, the Vaccination Reference Model will be analyzed in terms of its sensitivity to the non-pharmacological measures. These measures include actions like social distancing and maximum number of people in closed spaces, lock-downs, and mass testing. Secondly, the sensitivity of the model to the pharmacological measures, namely, the vaccination, will also be analyzed and discussed.

4.2.1 Sensitivity Analysis of Non-pharmacological Measures

The main purpose of these social interventions is to reduce the mentioned average reproductive number, R_e, previously introduced in Equation 2.7, which has to be

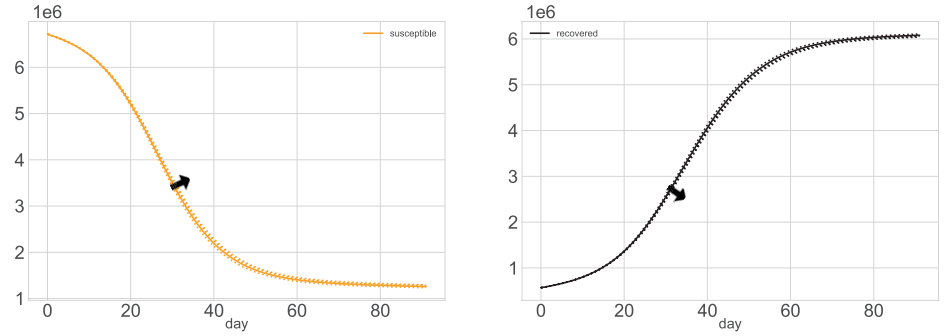
less than 1 for a pandemic to mathematically stop. These interventions are crucial, since understanding their impact and effect on the development of a pandemic will determine the effect of such actions on the social and economical costs they bring along. For example, all commerce and locals closed would have a big impact on the effective contact rate of the inhabitants of a population, but would also bring along a devastating economic aftermath. Therefore, there is a need to find an equilibrium to such contingency measures.

In this section, three different parameters are going to be modified in order to study the sensitivity of the Vaccination Reference Model to these parameters: The isolation rate of symptomatic individuals, κ_s, the isolation rate of asymptomatic individuals, κ_a, and lastly the replication factor, θ, which recalling from Section 2.2.1, can be expressed as $\theta = \gamma\ R_0$.

Sensitivity analysis of the model to κ_s

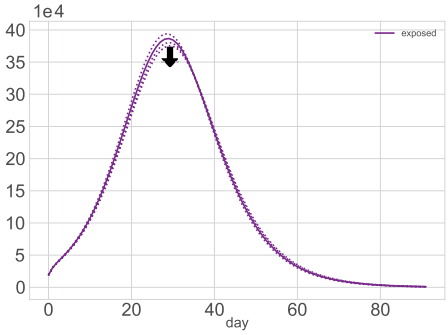
For this parametrization, κ_s has adopted the values of 0.7, 0.8 (reference), 0.9 and 1. The rest of the parameters remain as specified previously. There exist many types of detection tests for COVID-19, being the main ones the antigen test, PCR, and antibodies test, although this last one might be more useful in measuring the durability of vaccine responses. They all have different specifities and sensitivities, and also depend on the day of testing after infection. Generally, PCR testing has higher sensitivity and specificity than antigen testing, therefore it is more reliable. On the contrary, it is noticeably more expensive and results are not immediate, taking usually 24 hours.

In Figure 4.5 it can be observed the response of the SEIsIaQR model to the different values of κ_s, where the continuous line represents the Vaccination Reference Model, the dotted lines are the curves for the different values of such parameter, and the black arrows show the tendency of the dotted lines when augmenting the desired parameter. As it can be observed at first sight in Figures 4.5(a) and 4.5(b), there is not a drastic change on the susceptible and recovered proportion of the population. For Figures 4.5(c), 4.5(d) and 4.5(e) there is a slightly discernible tendency to flatten the curve, but the most noticeable change is logically in the infectious symptomatic curve 4.5(f). Here, a 10% increase in the value of κ_s (from 0.8 to 0.9) implies a descent of the curve by 11.02%, and, furthermore, increasing the value of this isolation rate for symptomatic individuals by 20% (meaning its value would be 1, which would be idealistic but unfeasible, as no detection test has an efficacy of 100%) would lead to a reduction of the curve by 19.91% from the initial representation. This implies that the flattening of the curve is not exponential, and the realistically attainable values of κ_s (0.8-0.9) do not suppose a drastic

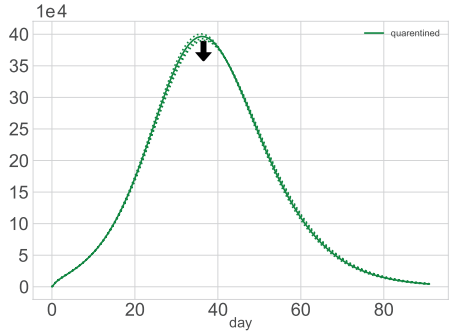


(a) Susceptible

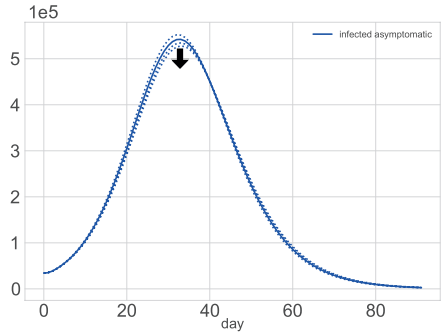
(b) Recovered



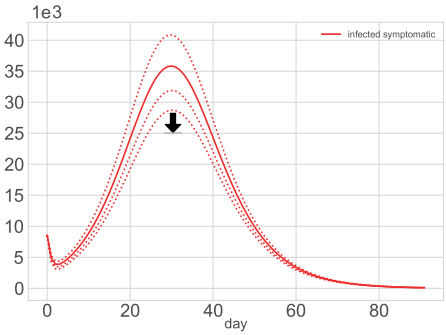
(c) Exposed



(d) Quarantined



(e) Infectious asymptomatic



(f) Infectious symptomatic

Figure 4.5: Sensitivity of the Vaccination Reference Model to the parameter κ_s.

change in the reduction of the amount of contagions by symptomatic individuals.

To match the previously represented correlations in Figure 4.5, the most effective changes in order to achieve those changes would be to use the best test in terms of sensitivity and specificity to detect COVID-19. Therefore, PCR testings would have to be applied to all those individuals who are presenting symptoms.

This carries along two obvious disadvantages: Firstly, and as it has been mentioned previously, this solution would exceed economical budgets for most governments, as the PCR is the priciest test to detect COVID-19. Secondly, and in relation to the first point, the results from lowering the curve by modifying κ_s is not as efficient (11.02% curve reduction given a 10% increase) for such an expenditure.

In the following, the results for the simulation including the modification of κ_s are included. Note that these results will contain the numeric solution to the reference simulation, which is carried out using the value $\kappa_s = 0.9$, for the sake of simplicity and further analysis.

	E	I_s	I_a	Q	**Day**
mean	126817.885041	10806.706145	190569.836347	149067.568407	45.50000
std	129404.881118	10771.574178	181604.113619	134081.913380	26.70206
min	1000.388903	94.781611	2924.798109	0.000000	0.00000
25%	13438.096028	1268.504849	34384.617792	28340.908360	22.75000
50%	72844.116342	6500.005839	118576.134043	101466.753552	45.50000
75%	227950.485276	19172.162672	341014.990731	263253.247341	68.25000
max	380565.517865	31902.864457	534386.375529	393075.416295	91.00000

Table 4.2: Sensitivity of the model assuming $k_s = 0.9$.

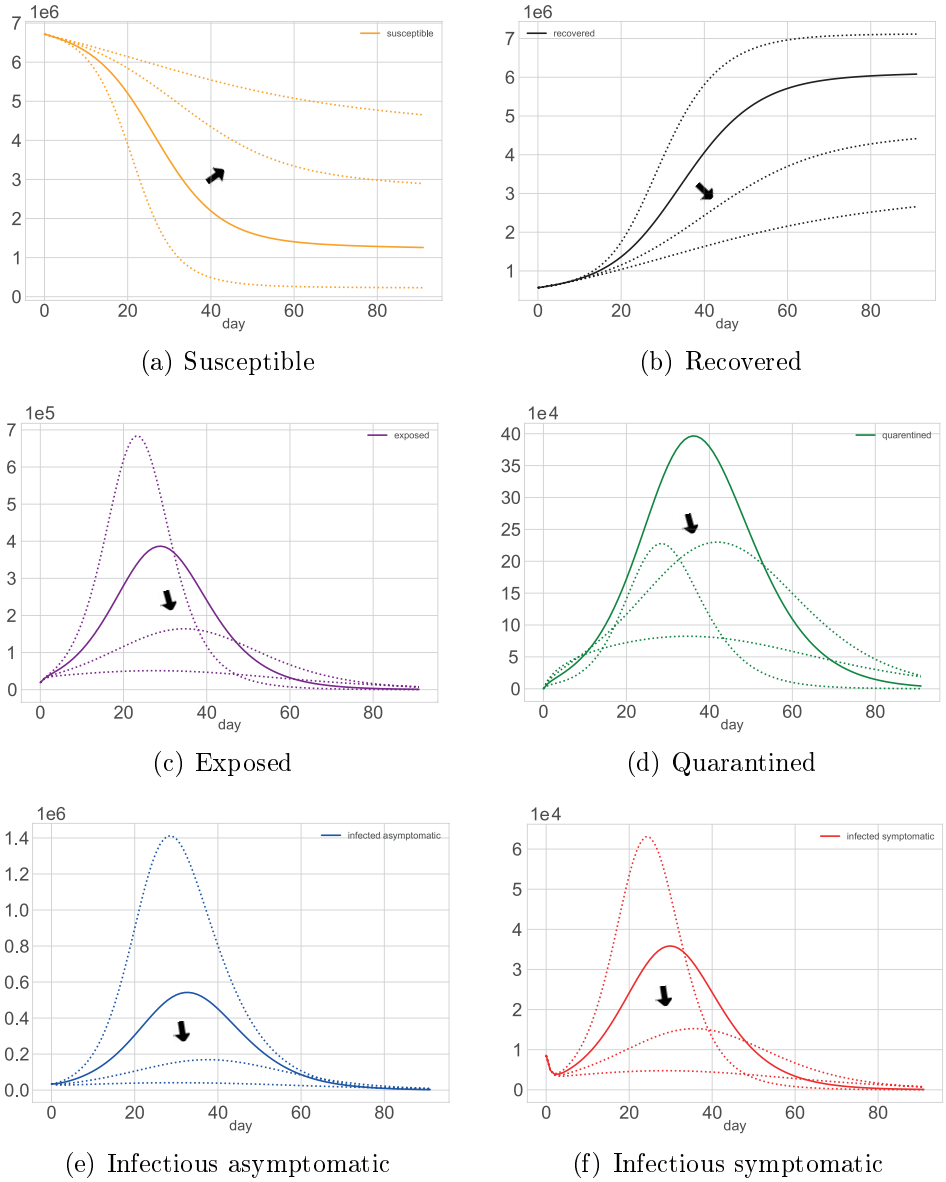
As observed in Table 4.2, there is not a drastic change in the number of affected individuals per compartment in comparison with the reference model. Therefore, we can conclude that the model is not very sensitive to changes in κ_s, and therefore it does not provide with an implicit solution to the reduction of the average reproductive number R_e.

Sensitivity analysis of the model to κ_a

In this case, κ_a will take the values of 0, 0.1 (reference), 0.2 and 0.3. This means that, for the reference model, only 10% of those who are asymptomatic are detected. This is a very difficult task, as they are thus less likely to seek out testing for the virus, unknowingly spreading the infection to other members of the community.

The SEIsIaQR model's response to the variation of κ_a is shown in Figure 4.6, being the continuous line the Vaccination Reference Model, and the dotted lines representing the parameter variation.

Contrary to κ_s, the model presents high sensitivity to κ_a. In Figures 4.6(a) and 4.6(b) it can be seen the faster rate of transition from the susceptible compartment to the recovered compartment as κ_a increases. For 4.6(c), 4.5(d), 4.5(e) and 4.6(f) there is an extremely perceptible tendency to flattening the curve, in opposition



(a) Susceptible (b) Recovered

(c) Exposed (d) Quarantined

(e) Infectious asymptomatic (f) Infectious symptomatic

Figure 4.6: Sensitivity of the Vaccination Reference Model to the parameter κ_a.

to the impact of κ_s in the previous case, where mainly the infectious symptomatic curve was affected. Hence, a 10% increase in the value of κ_a (from 0.1 to 0.2) produces a flattening of 57.45% of the curve of infected symptomatic individuals. Moreover, increasing the value of this parameter by 20% (from 0.1 to 0.3) would lead to an almost complete flattening of the curve, reducing it by 76.07% from the initial representation, and being the maximum number of infectious symptomatic

individuals the initial value of 8578. Although, in the case of 4.6(d), it can be observed a more complex behavior of the curve. A κ_a value of 0.1 (reference) involves more quarantined individuals than a κ_a value of 0, but more than a value of 0.2. This can be explained as, when κ_a is 0, the detection and acknowledgement of asymptomatic individuals is more complicated than with a detection rate of 0.1. on the other hand, when the detection rate is 0.2, there will be less susceptible and less exposed individuals, lowering the quarantined curve. All of this can be also observed in Table 4.3 for a value of κ_a of 0.2:

	E	I_s	I_a	Q	**Day**
mean	81780.512569	7755.447932	86261.990665	120418.319113	45.50000
std	53066.704946	4831.152416	53189.081893	72426.809141	26.70206
min	6755.112633	689.434441	9285.241276	0.000000	0.00000
25%	33721.490988	3691.990307	36973.917719	52988.913062	22.75000
50%	75510.578810	7264.319691	79444.401575	114000.822174	45.50000
75%	132172.572130	12288.248130	137197.485682	189476.923668	68.25000
max	164022.731198	15253.944220	168761.439599	230078.389729	91.00000

Table 4.3: Sensitivity of the model assuming $k_a = 0.2$.

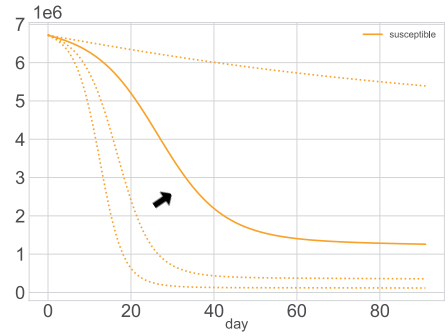
Therefore, the model is indeed very sensitive to changes in the parameter κ_a, as slight variations in this number imply dramatic outcomes. Hence, according to the model, an isolation rate of 30% of the asymptomatic individuals would drastically reduce the rate of infections. This outcome can be achieved with measures like massive testing of the population. Conversely to the previous section, in which κ_s is analyzed, this measure would not suppose an utmost economic expenditure, as the rapid testing would be more dynamic, cheap, and efficient enough for the purpose of incrementing 10-20% the parameter κ_a.

Sensitivity analysis of the model to θ

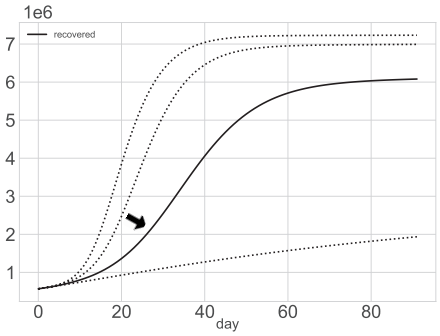
Finally, the last non-pharmacological parameter to be analyzed will be the replication factor θ. Recalling from Section 2.2.1, θ is the product of the interaction factor γ and R_0. Therefore, the analysis of the sensitivity of the model to this parameter will depend on the variation of the interaction factor γ. Hence, γ will be 1, 2 (reference), 3 and 4. In this case, higher values of γ will signify less strict social distancing and social interaction measures. On the contrary, a value of 1 symbolizes no effective contact among individuals, hence not being possible the spread of the virus.

A value of 2 for the interaction factor γ in the Reference Vaccine Model has been chosen due to the current social distancing and restriction measures at the time of

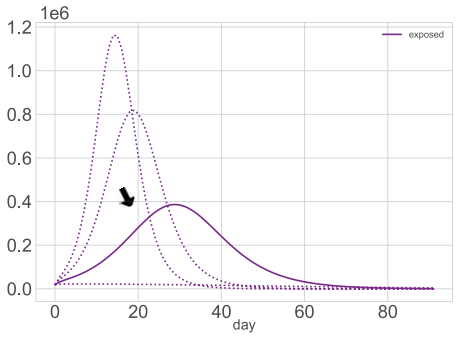
writing this report, such as the closure of locals and the curfew (23:00 in Madrid), the maximum number of people allowed for a gathering (maximum of 4), and so on.



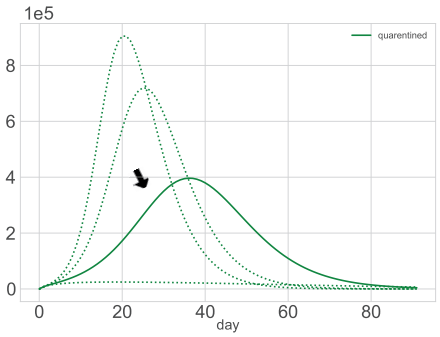
(a) Susceptible



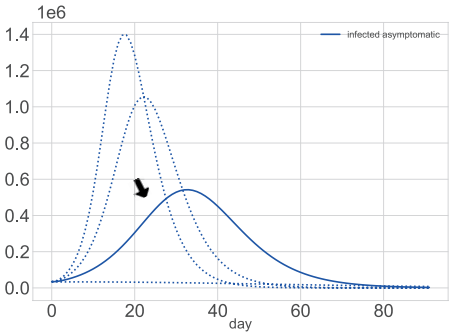
(b) Recovered



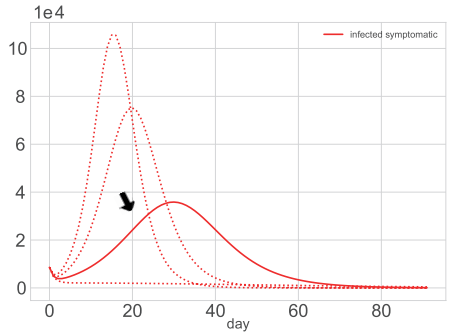
(c) Exposed



(d) Quarantined



(e) Infectious asymptomatic



(f) Infectious symptomatic

Figure 4.7: Sensitivity of the Vaccination Reference Model to the parameter θ.

Once again, there is a tendency to flatten all curves as shown in Figure 4.7, where a value of 1 to the interaction factor would mean an evident stop of the spread of the virus. On the contrary, if there was a moderation of the social inter-

action measures (a rise in the value of γ to 3) at day 0, for example, by delaying the curfew a couple of hours or allowing a larger amount of people in a gathering, there would be a significant change in the maximum of infected symptomatic individuals, going from 35853 at day 30, to a much quicker rate of infection, reaching 75159 by day 20, an increase of 109.63%. It can be seen in Table 4.4 how the maxima are significantly lower for a θ value of 3:

	E	I_s	I_a	Q	**Day**
mean	118760.219767	11207.464053	178573.547017	138933.633124	45.50000
std	120216.323749	11067.220152	168745.820202	124013.334381	26.70206
min	861.780288	92.071627	2585.072859	0.000000	0.00000
25%	12302.472836	1304.438715	33634.394525	26373.116092	22.75000
50%	69685.847191	6910.152989	113686.488842	96447.784166	45.50000
75%	213129.272659	19816.413841	317987.562822	248711.109075	68.25000
max	351874.140703	32641.189143	494812.421540	363071.612177	91.00000

Table 4.4: Sensitivity of the model assuming $\theta = 3$.

Note that for Table 4.4, a value of 3 has been chosen, even if the reference model includes a θ value of 2, and thus, the next logic step would be to set it to 1, but as discussed previously, a value of 1 would imply a total strict lock down, with a consequent complete flattening of the curve. Therefore, for a more realistic approach, the analysis includes the improvements of the reduction of the replication factor θ from 3 to 2.

To conclude, the model is very sensitive to variations in γ, as seen in Figure 4.7. Thus, it can be stated that both θ and κ_a have a bigger impact on the outcome of the model. Therefore, it may be affirmed that, with the right studies and professionals, and the correct combination of measures, a pandemic can be controlled in the long term with the most efficient use of resources.

4.2.2 Sensitivity Analysis of Pharmacological Measures

Lastly, it is important to analyze the sensitivity of the model to the variation of the vaccination rate v. This rate has been assumed to be 0.0015 [12], meaning that, approximately, 0.15% of the total susceptible population will receive a dose daily. The increase in this rate requires a major economic, logistic and administrative expense. Although, it will be shown how theoretically this vaccination rate plays an important role in the control of COVID-19.

The values of v chosen for the sensitivity analysis of the model are following an exponential growth for the sake of the visualization, being 0.0015 (reference), 0.003,

0.006 and 0.012. As mentioned, the fact of achieving a greater vaccine administration rate is dependent of some other factors. Firstly, the required doses to be administered have to be available, as vaccine acquisition and partitioning among countries has been a difficult task. Secondly, it is crucial to have an efficient logistic system, which also protects the vaccines and all their requirements for safe transportation. Lastly, there is a need of well-organized and methodical staff to coordinate the administration of the vaccine to comply with the pertinent safety measures. As seen in Figure 4.8, the acceleration in the vaccination rate produces an exponential flattening of the curves. Also, it can be seen in Figure 4.8(a) how the tendency of recovered individuals curve would be more linear, and in consequence, reducing as well the amount of exposed people, as seen in Figure 4.8(c). In the first place, doubling the vaccination rate, which would imply doubling the acquisition of such vaccines and their weekly application, would produce a flattening of the curve of infected symptomatic individuals of 8.96%, and a vaccination rate of 0.006, a reduction of 24.65%. Below in Table 4.5, the numerical results for $v = 0.003$ can be observed:

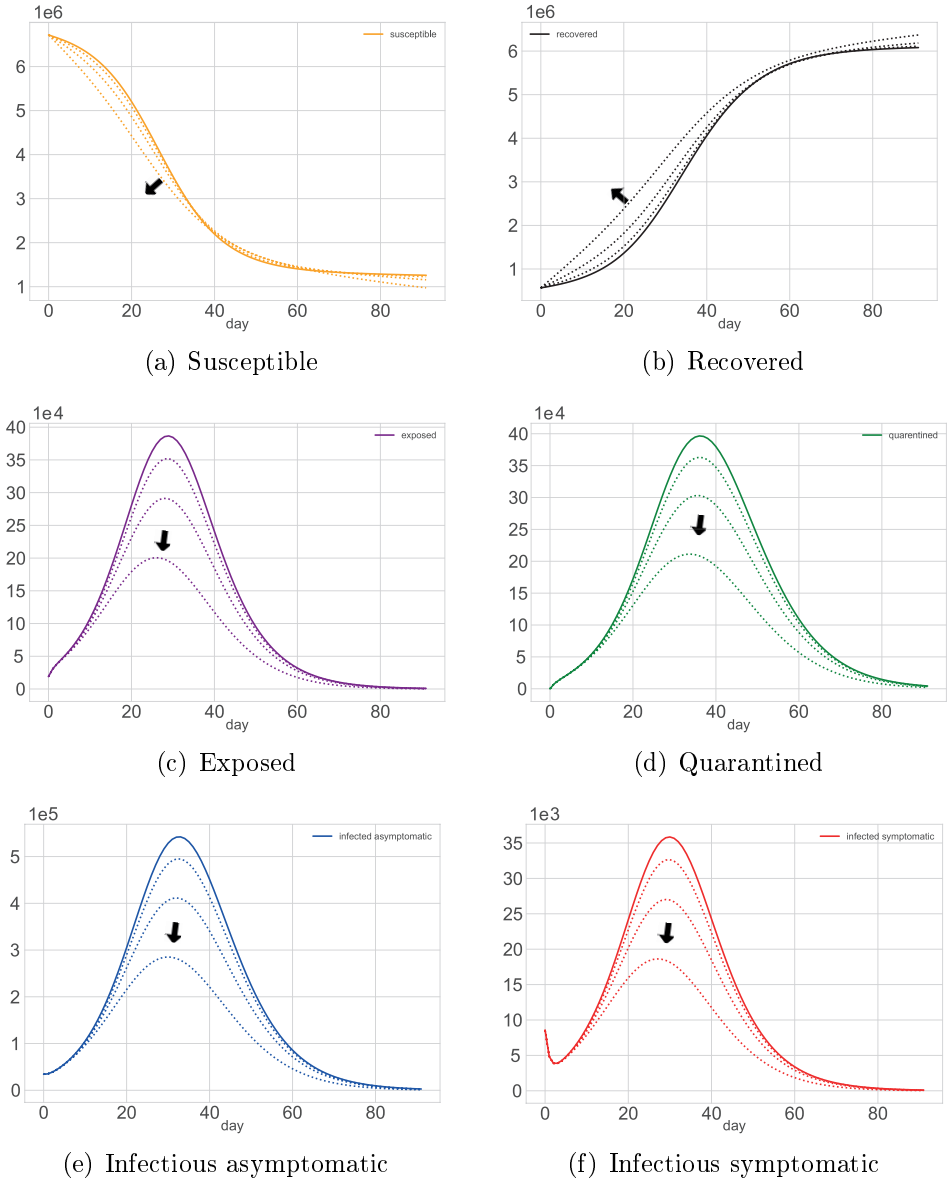
	E	I_s	I_a	Q	**Day**
mean	154493.131403	14538.586715	2.319681e+05	180773.712022	45.50000
std	247666.232777	22822.026091	3.340756e+05	236415.399320	26.70206
min	4.864492	0.557277	2.813660e+01	0.000000	0.00000
25%	253.257724	28.960618	1.423659e+03	3936.622695	22.75000
50%	13279.345194	1532.048824	4.419875e+04	49370.136099	45.50000
75%	199543.152220	19032.210334	3.535655e+05	303601.998343	68.25000
max	818190.243187	75159.566235	1.055891e+06	719647.205463	91.00000

Table 4.5: Sensitivity of the model assuming $v = 0.003$.

Accordingly, it can be observed that vaccination, at these rates, might not be the most effective measure in the short term compared to others analyzed previously, like mass testing or social distancing. However, vaccination is the key for a long term stoppage of the spread of the virus, as vaccinated individuals are permanently removed from the susceptible compartment, whereas those individuals who have been infected and not vaccinated could have antibodies up to the eighth month after infection, which in any case prevents contagions in the period of time prior to the eighth month.

In Table 4.6, the results for the simulation of the variables E, I_s, I_a, and Q have been displayed for a better visualization of the overall impact of the modification of the parameters in the equations.

As seen, the most noticeable remarks come down to the sensitivity of the model to the variations of the isolation of asymptomatic individuals k_a, as well as the inter-



(a) Susceptible (b) Recovered

(c) Exposed (d) Quarantined

(e) Infectious asymptomatic (f) Infectious symptomatic

Figure 4.8: Sensitivity of the Vaccination Reference Model to the parameter v.

action factor θ. Moreover, the most significant increments are given by k_a taking into account the variation of the parameter itself, as the value of this parameter goes from 0.1 to 0.2 ($\Delta 10\%$), whereas θ, which variations are also very sensitive to the model, suffers a greater variation in order to explore such results.

On the other hand, it can be seen that variations in the isolation rate for symp-

Parameter	$\Delta\%$	$\Delta E(\%)$	$\Delta Q(\%)$	$\Delta I_a(\%)$	$\Delta I_s(\%)$
k_s	10	-1.57	-0.88	-11.02	-1.43
k_a	10	-57.57	-41.98	-68.87	-57.45
θ	50	-111.61	-81.47	-94.77	-109.63
v	100	-9	-8.45	-8.73	-8.96

Table 4.6: Summary of the parametric simulations.

tomatic individuals k_s do not provide with very valuable information to the model in comparison to the rest of the parameters. Moreover, the vaccination rate v also does not seem to supply further knowledge when it comes to the sensitivity of the model to its variations. Nevertheless, the vaccination rate considered (0.0015 in the reference model and 0.003 for the first simulation) is at very low values due to the maturity of the vaccination programs during the time of study. Despite that, vaccination constitutes a really effective measure against the COVID-19 pandemic, as the vaccinated individuals are no longer exposed and directly moved to the recovered compartment.

4.3 Sensitivity analysis: Stochastic Approach

In the previous experiments, it has been assumed that both the pharmacological and non-pharmacological measures are represented as deterministic values, and therefore, the changes in the system overtime have been evaluated by adjusting its different parameters. For the following part of the numerical experimentation, a further step has been taken for more validity: The application of the Monte Carlo method to evaluate the behavior of the model assuming that the mitigation measures are represented by random variables, as we cannot know for certain what the exact values for these parameters are, providing a more realistic insight. For this, it is presumed that the values for such parameters fall within a certain range given by a random distribution.

The Monte Carlo Simulation, or Monte Carlo Method [17], is a mathematical technique, which is used to estimate the possible outcomes of an uncertain event, by assigning multiple values to uncertain variables to achieve multiple results and then to average the results to obtain an estimate. Thus, its goal is to predict a set of outcomes based on an estimated range of values versus a set of fixed input values. Therefore, these simulations run thousands of times to determine the statistical range of values for the predicted outcomes. As the number of simulations increase, the number of forecasts also does, allowing to project outcomes with more accuracy.

This is especially useful to study the evolution of the pandemic in different scenarios, as the measures are not going to be always the same for every city, and even for every district within a city. Moreover, the behavior of each individual with respect to the pertinent restrictions does not allow for a constant parametrization of such measures. Therefore, this can be represented by the assumption of these parameters as random variables.

For this part of the experimentation, the vaccination rate (pharmacological measure) has been assumed to be constant, but the non-pharmacological parameters have been hypothesized to be random variables:

- The replication factor θ has been set to be a gamma distribution G(4200, 0.001) centered at 4.2 (recall θ as the product between R_0 and the interaction factor γ, being 2 the value for γ in the reference model).
- κ_s is a beta distribution B(35.55, 8.88) centered at 0.8.
- Lastly, κ_a will also be a beta distribution B(10, 90), which will be centered at 0.1.

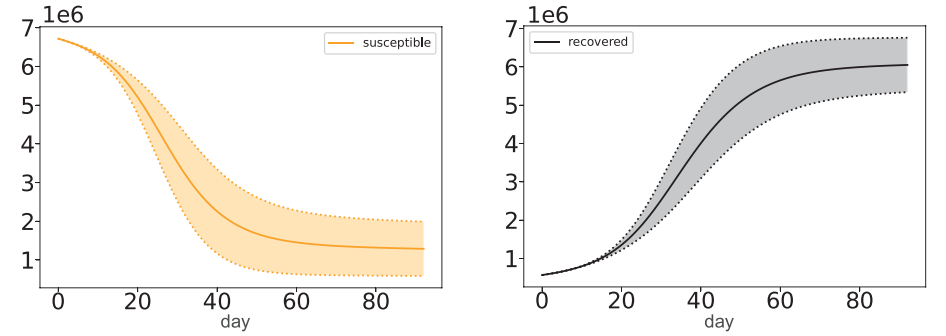
As seen, the centers of the distributions have been chosen according to the values of the reference model in the parametric analysis. Moreover, for each distribution a set of 10000 samples has been generated. Therefore, the algorithm will perform 10000 simulations, selecting values at random for θ, k_s, and k_a from each of the specified distributions. This will ensure more statistical validity. Lastly, 95% confidence intervals have been calculated in order to graphically represent the fore-mentioned envelopes.

As it can be observed in Figure 4.9, the combined variability given to the model by each of the three random parameters (θ, k_s and k_a) causes the envelopes (95% confidence intervals) to be wider, specially towards the apex of the curves.

Curve	**Mean**	**CI High**	**CI Low**	**Day**
Exposed	384704	540814	228595	28
Quarantined	379489	447548	311430	36
Infected asymptomatic	557366	879378	235353	32
Infected symptomatic	35843	51331	20355	29

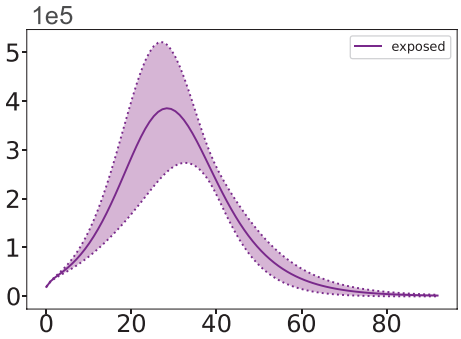
Table 4.7: Monte Carlo simulation results.

These confidence intervals are more uncertain and unclear for curves like Figure 4.9(e), where the range goes from 270612 to 628463 by day 40. On the other hand,

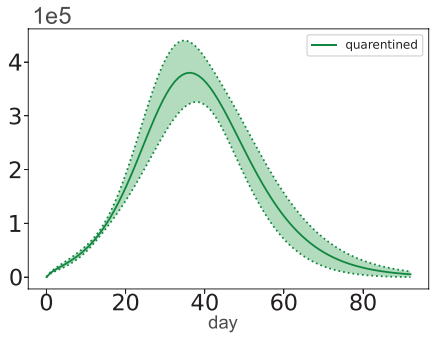


(a) Susceptible

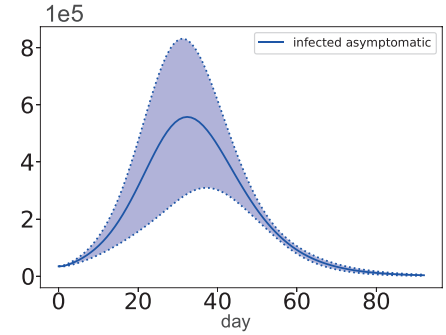
(b) Recovered



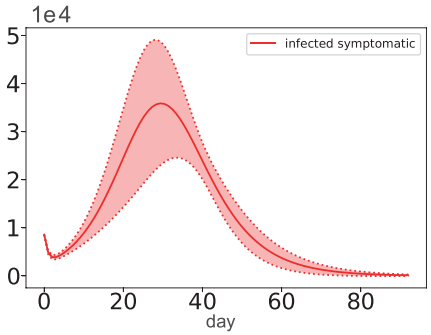
(c) Exposed



(d) Quarantined



(e) Infectious asymptomatic



(f) Infectious symptomatic

Figure 4.9: Mean values and 95% confidence interval envelopes.

these intervals provide more valuable information for curves like Figure 4.9(f), where it can be stated that, by day 40, there will be a 95% chance that the mean number of infected asymptomatic individuals lies between 18840 and 27539. For a more detailed look, Table 4.7 displays the mean and confidence interval for each compartment.

As observed, the peak of each curve lies in a close range interval of days, being the maximum 8 days between the exposed and quarantined individuals. Therefore, the evolution of these curves are closely related to one another, and it can be an advantage in order to establish contingency measures and evaluating them as such, as a time reference is very valuable. Finally, the 95% has defined a wider spectrum for the enveloped of the curves, where the quarantined stands out to be the most precise one, as its confidence intervals are closer to the mean value curve.

Chapter 5

Conclusions

5.1 Acquired Skills

One of the main reasons for the selection of this project was the feasibility of application of the knowledge developed during the past four years, specially those regarding differential equations and *Python* programming. Moreover, the fact that this project tries to solve an actual real-life problem makes it more challenging and attractive. This being said, the main competences that had to be gained include:

- More extensive knowledge to add on to the previously obtained during the years regarding differential equations, including the understanding, parameter modification and solving of systems of non-linear differential equations.
- Consequently, the understanding of the SIR model has been essential for the further development of the project. Moreover, the comprehension of these compartmental systems, along with the previous point, allow for an improved SIR model, the modified SEIQR model to be able to fit our specific problem, with each one of its characteristics.
- Finally, and as a result of the previous two points, the implementation of all the information obtained into the *Python* software in an easy, understandable and replicable way, thank to the function *odeint*, which solves the initial value problem for stiff or non-stiff systems of ordinary differential equations, using the *lsoda* function from the FORTRAN library *odepack*.

5.2 Final Conclusions

Regarding the conclusions of the project itself, and after carefully analyzing the information and results provided by *Chapter 2*, *Chapter 3*, *Chapter 4*, and after observing the development of the pandemic situation throughout the months, it can be stated that:

- First and foremost, and being one of the main objectives, the Vaccination Reference Model used for the development of the project, as well as the outcomes of the different experimentation procedures for the respective months of study, come close to reality, making the model scenarios realistic, and thus, giving the model more predictive power. This highlights the importance of this sort of studies and professionals like mathematicians and epidemiologists in the task of analyzing and controlling pandemic grows.

- It is key to understand, after reading *Chapter 4* and *Chapter 5*, that not all measures are equally effective during the development of a pandemic. This is possible to observe thank to the sensitivity analysis of the different parameters relevant for our study:

  - The most theoretically effective measures are the isolation rate of asymptomatic individuals, possible through initiatives like mass-testing and the interaction factor, which can be modeled by the degree of freedom regarding social interaction, from reducing the maximum capacity in public spaces, determining a curfew, or even a complete lock-down. Regarding this, some measures are more economically aggressive than others, so it is crucial to correctly analyze and study the situation to find the balance between health and wealth.
  - On the other hand, the model has shown less sensitivity to the isolation rate of symptomatic individuals. This can be due to the already high number established for this parameter, namely 0.8, meaning that almost all individuals who present symptoms are detected and placed in quarantine.

  Moreover, the vaccination rate, although it has not shown as much sensitivity as the isolation rate of asymptomatic individuals or the interaction factor, is essential for the stoppage of the virus spread, as it places susceptible individuals who have been vaccinated directly into the recovered compartment, which accelerates the ride-out process of the pandemic.

- Complementing the parametric analysis with a stochastic study including Monte Carlo simulations adds more valuable information, as the real-life situations are not deterministic and are influenced by a certain degree of randomness. This allows for a more pragmatic approach of the problem, and therefore, providing a feasible set of results based on the randomized allocation of the model parameters. Hence, the predictive power of the model is improved by combining the knowledge provided by both the parametric and the stochastic analysis.

- Finally, it has been learned that compartmental models based on differential equations, like the one proposed in this project, can be applied to any type of epidemic outbreak, since the model, the equations and the parameters can be modified to fit any situation. Therefore, this makes itself a possible future line of research, as it has been demonstrated the importance of this type of studies during a pandemic situation.

5.3 Future Lines of Research

The research and experimentation proposed during this project could be considered valid for the fundamentals of a predictive decision-making model, but could be further improved to achieve a higher level of accuracy:

- In the first place, the parameters used for our model have been assumed from previous descriptive studies, but for a more solid and accurate model, a study to establish the values of such parameters, adapting it to our target population, could be performed, requiring more time and effort. This involves advanced knowledge in statistics and inference, with more complex techniques for parameter estimation like Bayesian inference.

- Secondly, the population set-up has been an evolving field, and so have been the parameters of the model. For instance, the values for the social distancing measures represented by θ have been decreasing overtime, producing as well variations in the model. Therefore, for a more realistic approach, the values of θ could have been set to decrease through the course of time.

- In relation to the previous point, the vaccination rate v has been also assumed to be constant and at a very low rate, which in reality has been exponentially growing as more doses of the vaccine where arriving to the country. Thus, v could also be set to grow overtime.

- Moreover, the efficacy of the vaccines have also been assumed constant in the computation of v, namely 0.95, but in reality there have been, at least, 3

different types of vaccines during the months of study (*Pizer*, *Moderna* and *AstraZeneca*), each of them with different levels of efficacy. Furthermore, the vaccination dosage levels of efficacy are also different among vaccines, having a high level of protection after one dose, and the highest after the second one.

- Lastly, v could also have been treated as a random variable for the stochastic study, as the vaccination rate has been slightly different depending on the population area.

Chapter 6

Annexes

6.1 Annex A: Parametric Analysis Code

```
# import libraries
import pandas as pd
import numpy as np
from scipy.integrate import odeint
import matplotlib.pyplot as plt

# the SEIQR model differential equations.
def deriv1(state, t, N, theta, beta, alpha, ks, ka, Tinf, Tser, Tlat, v):
S, E, Is, Ia, Q, R= state
# Change in S population over time
dSdt = -theta/Tinf * (Is + alpha*Ia)*S/N - v*S
dEdt = theta/Tinf * (Is + alpha*Ia)*S/N - E/Tlat
dIsdt = (1-beta)*E/Tlat - (ks + 1/Tinf)*Is
dIadt = beta*E/Tlat - (ka + 1/Tinf)*Ia
dQdt = ks*Is + ka*Ia - Q/Tser
dRdt = (Is+Ia)/Tinf + Q/Tser + v*S
dCdt = (1 - beta) * E/Tlat + ka*Ia
return dSdt, dEdt, dIsdt, dIadt, dQdt, dRdt

# parametrization (parameters can be modified to analyze sensitivity)
effective_contact_rate = 0.3
recovery_rate = 1/7
interaction_factor = 2
# R0
print("R0 is", effective_contact_rate / recovery_rate)
```

```
# theta
theta = effective_contact_rate / recovery_rate * interaction_factor
# ks
isolation_rate_s = 0.8
# ka
isolation_rate_a = 0.1
# v
vaccine = 0.0015 * 0.95
# beta
asymptomatic_population_ratio = 0.8
#alpha
infectiousness_ratio = 1
# Tser
serial_interval = 7.5
# Tlat
mean_incubation_period = 5.2
# Tinf
Infectious_period = serial_interval - mean_incubation_period

# target population set up
total_pop = 6779888
exposed = 18963
recovered = 586913-exposed
quarentined = 0
infectious_s = 8578
infectious_a = asymptomatic_population_ratio * infectious_s / (1-asymptomatic_populati
susceptible = total_pop - exposed - infectious_s/(1-asymptomatic_population_ratio)
# days for study
days = range(0, 92)
# differential equations on target population
ret = odeint(deriv1,
[susceptible, exposed, infectious_s, infectious_a, quarentined, recovered] ,
days,
args=(total_pop, theta, asymptomatic_population_ratio, infectiousness_ratio,
isolation_rate_s, isolation_rate_a, serial_interval, mean_incubation_period,
Infectious_period, vaccine))
S, E, Is, Ia, Q, R = ret.T

# build dataframe
dfv = pd.DataFrame({
```

```
'susceptible': S,
'exposed' : E,
'infected symptomatic': Is,
'infected asymptomatic': Ia,
'quarentined': Q,
'recovered': R,
'day': days})

# plot, where y can be ['susceptible', 'recovered', 'exposed', 'quarantined', 'infected asymptomatic', 'infected symptomatic']
dfv.plot(x='day',
y=['infected symptomatic'],
color=['red'],
stacked=False)
```

6.2 Annex B: Stochastic Analysis Code

```
# import libraries
import pandas as pd
import numpy as np
from scipy.integrate import odeint
import matplotlib.pyplot as plt
import scipy.stats
import statistics
import random

# the SEIQR model differential equations.
def deriv1(state, t, N, theta, beta, alpha, ks, ka, Tinf, Tser, Tlat, v):
S, E, Is, Ia, Q, R= state
# Change in S population over time
dSdt = -theta/Tinf * (Is + alpha*Ia)*S/N - v*S
dEdt = theta/Tinf * (Is + alpha*Ia)*S/N - E/Tlat
dIsdt = (1-beta)*E/Tlat - (ks + 1/Tinf)*Is
dIadt = beta*E/Tlat - (ka + 1/Tinf)*Ia
dQdt = ks*Is + ka*Ia - Q/Tser
dRdt = (Is+Ia)/Tinf + Q/Tser + v*S
dCdt = (1 - beta) * E/Tlat + ka*Ia
return dSdt, dEdt, dIsdt, dIadt, dQdt, dRdt

# create samples
# theta
sample1 = np.random.gamma(4200, 0.001, 10000)
# ks
sample2 = np.random.beta(35.555, 8.888, 10000)
# ka
sample3 = np.random.beta(10, 90, 10000)

# Monte Carlo simulations
dct = { }
for i in range(0, 10000):
effective_contact_rate = 0.3
recovery_rate = 1/7
interaction_factor = 2
# R0
print("R0 is", effective_contact_rate / recovery_rate)
# theta
```

```
theta = sample1[i]
# ks
isolation_rate_s = sample2[i]
# ka
isolation_rate_a = sample3[i]
# v
vaccine = 0.0015 * 0.95
# beta
asymptomatic_population_ratio = 0.8
#alpha
infectiousness_ratio = 1
# Tser
serial_interval = 7.5
# Tlat
mean_incubation_period = 5.2
# Tinf
Infectious_period = serial_interval - mean_incubation_period

# target population set up
total_pop = 6779888
exposed = 18963
recovered = 586913-exposed
quarentined = 0
infectious_s = 8578
infectious_a = asymptomatic_population_ratio * infectious_s / (1-asymptomatic_populati
susceptible = total_pop - exposed - infectious_s/(1-asymptomatic_population_ratio)
# days for study
days = range(0, 92)
# differential equations on target population
ret = odeint(deriv1,
[susceptible, exposed, infectious_s, infectious_a, quarentined, recovered] ,
days,
args=(total_pop, theta, asymptomatic_population_ratio, infectiousness_ratio,
isolation_rate_s, isolation_rate_a, serial_interval, mean_incubation_period,
Infectious_period, vaccine))
S, E, Is, Ia, Q, R = ret.T

# build dataframe
dct['df%s' % i] = pd.DataFrame({
'susceptible': S,
```

```
'exposed' : E,
'infected symptomatic': Is,
'infected asymptomatic': Ia,
'quarentined': Q,
'recovered': R,
'day': days})

# mean and confidence intervals
df_concat = pd.concat(dct
df1=df_concat.groupby('day', as_index=True).mean()
df2=df_concat.groupby('day', as_index=True).std()
df3= df1+ (1.645*df2)
df4= df1- (1.645*df2)

# plot, where y, y2, y3 can be ['susceptible', 'recovered', 'exposed', 'quarantined',
'infected asymptomatic', 'infected symptomatic']
x = np.linspace(0,92,92)
y = df1['susceptible']
y2 = df3['susceptible']
y3 = df4['susceptible']
plt.plot(x, y, color='orange', label='susceptible')
plt.legend()
plt.plot(x, y2, color='orange', linestyle='dotted')
plt.plot(x, y3, color='orange', linestyle='dotted')
plt.fill_between(x, y2, y3, color='orange', alpha=0.25)
```

Bibliography

[1] World Health Organization *Coronavirus*, Webpage: https://www.who.int/health-topics/coronavirus#tab=tab_1.

[2] Steven C. Chapra and Raymond P. Canale, *Numerical Methods for engineers*, sixth edition, McGraw-Hill, 2010.

[3] H. Weiss *The SIR model and the foundations of public health*, Materials Matemàtics, 2013.

[4] Qun Li, et al., *Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia*, N. Engl. J. Med., 2020.

[5] Joseph T. Wu, Kathy Leung, and Gabriel M. Leung, *Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study*, The Lancet 395, 2020.

[6] Alberto Olivares and Ernesto Staffetti, *Uncertainty quantification of a mathematical model of COVID-19 transmission dynamics with mass vaccination strategy*, Chaos, Solitons, and Fractals, Elsevier Ltd., May 2021.

[7] Eduardo L. Brugnago, Rafael M. da Silva , Cesar Manchein, and Marcus W. Beims, *How relevant is the decision of containment measures against COVID-19 applied ahead of time?*, Chaos Solitons Fractals, 2020.

[8] *Anuario Estadístico De La Comunidad De Madrid. Población y Hogares*, Webpage: https://www.madrid.org/iestadis/fijas/estructu/general/anuario/ianucap02.htm, Community of Madrid, 2021.

[9] Spanish Healthcare Ministry, Spanish Government *Actualización n^0 322. Enfermedad por el coronavirus (COVID-19). 01.03.2021*, Webpage: https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/documentos/Actualizacion_322_COVID-19.pdf, March 1_{st} 2021.

[10] Gaurav Goswami, Jayanti Prasad, and Mansi Dhuria, *Extracting the effective contact rate of COVID-19 pandemic*, School of Engineering and Applied Science, Ahmedabad University, India, 2020.

[11] Andrea L. Bertozzi, Elisa Franco, George Mohler, Martin B. Short, and Daniel Sledge, *The challenges of modeling and forecasting the spread of COVID-19*, Proceedings of the National Academy of Sciences of the United States of America, July 21_{st}, 2020.

[12] Spanish Healthcare Ministry, Spanish Government *Vaccination plan*, Webpage: `https://www.mscbs.gob.es/en/profesionales/saludPublica/prevPromocion/vacunaciones/covid19/vacunasCovid19.htm`.

[13] Kathy Katella, *Comparing the COVID-19 Vaccines: How Are They Different?*, Webpage: `https://www.yalemedicine.org/news/covid-19-vaccine-comparison`, June 23_{rd}, 2021.

[14] *Scipy.integrate.odeint*, Scipy.integrate.odeint - SciPy v1.7.0 Manual, Webpage: `https://docs.scipy.org/doc/scipy/reference/generated/scipy.integrate.odeint.html`.

[15] Spanish Healthcare Ministry, Spanish Government *Actualización n^0 344. Enfermedad por el coronavirus (COVID-19). 01.04.2021*, Webpage: `https://www.mscbs.gob.es/profesionales/saludPublica/ccayes/alertasActual/nCov/documentos/Actualizacion_344_COVID-19.pdf`, April 1_{st} 2021.

[16] Lauren M. Kucirka, Stephen A. Lauer, Oliver Laeyendecker, Denali Boon, and Justin Lessler, *Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure*, Ann Intern Med, August 18_{th}, 2020.

[17] Shlomo Mark and Shaul Mordechai, *Applications of Monte Carlo Methods in Science and Engineering*, INTECH Open Access Publisher, 2011.

[18] Hamdy Youssef, Najat Alghamdi, Magdy A. Ezzat, Alaa A. El-Bary, and Ahmed M. Shawky, *Study on the SEIQR model and applying the epidemiological rates of COVID-19 epidemic spread in Saudi Arabia*, Infect. Dis. Model, 2021.