MASTER'S THESIS

Master of Science in Advanced Mathematics and Mathematical Engineering

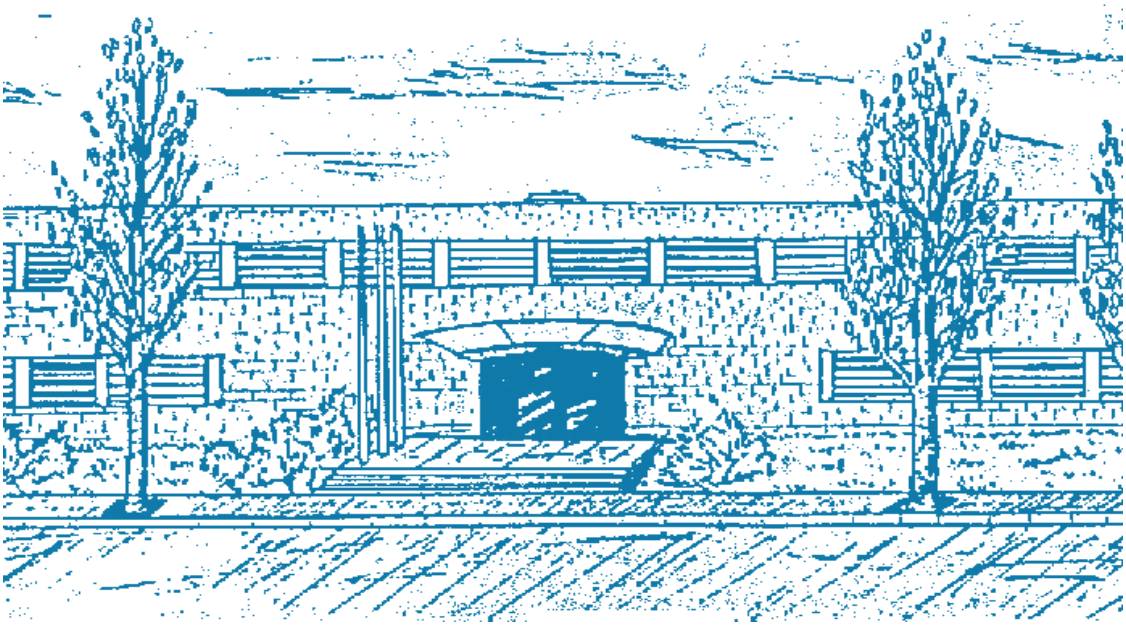
Title: Parameter estimation for a mathematical model predicting the COVID-19 spread in the *Àrea Metropolitana de Barcelona*

Author: Alicja Barbara Kubik

Advisors: Benjamín P. Ivorra, Ángel M. Ramos (Universidad Complutense de Madrid) and J. Tomás Lázaro (Universitat Politècnica de Catalunya)

Department: Departament de Matemàtiques, Facultat de Matemàtiques i Estadística

Academic year: 2019/2020



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Acknowledgments

In first place, I want to thank profs. Ángel Ramos and Benjamín Ivorra for accepting my proposal for tutoring me in this project, despite the fact they work in Madrid. Both have been on the front line of the mathematical modeling for the COVID-19 and it has been a honour learning from them in this my biggest first approach to mathematical modeling in epidemiology. It will be a great pleasure if I finally have the opportunity to work with both of you in few months.

I also thank prof. Tomás Lázaro for accepting being my tutor in the UPC. We have learned together through this adventure, but he also has been really helpful with the theoretical part and, more importantly, making me feel more confident about this project, since he has been encouraging me from the first moment we started to work together. If our paths decide to cross in the future, I will be pleased to collaborate again.

Finally, thanks to my parents and friends for supporting me during the project and during the whole master; it has not been easy reaching the end, but it could have been harder if they hadn't been there with me.

Abstract

Epidemics have been emerging and reemerging for thousands of years, killing millions of people, since the first reported plagues in the Bible until the current coronavirus, which has put in quarantine almost the whole world. Epidemiology, the science studying these diseases, has been being developed from Hippocrates' first medical studies until now, and it was not until the 18^{th} century with Daniel Bernouilli that there appeared a new branch: mathematical epidemiology.

During the following pages, we will first establish some basic concepts and a motivational introduction in order to present some common results on mathematical epidemiology. For that, it will be introduced some theory using the basic SIRS model so as to understand well the involved tools, such as Lyapunov theory on equilibrium points and their stability or the basic reproduction number R_0, which will be crucial to determine the future of the disease. Finally, we will try to apply all the learned concepts to study a more complex mathematical model, still in development, whose aim is trying to reproduce the behaviour of the well-known COVID-19; concretely, there will be shown some numerical results for its evolution in the Metropolitan Area of Barcelona.

Keywords: COVID-19, mathematical epidemiology, human infectious disease, disease spread, basic reproduction number, effective reproduction number, SIRS model, Lyapunov stability, θ-SEIHRD model, numerical simulation, parameter calibration.

Contents

1 Introduction

The word *epidemiology* comes from the Greek decomposition *epi* (upon), *demos* (the people) and *logos* (study, science), i.e., it literally means "the study of what is upon the people". More precisely, it is a scientific discipline inside the medical field in charge of studying the dynamics of a disease in a given population by understanding the causes of the disease, trying to predict its course and finally developing some ways to control it; although it is referred to any kind of disease (e.g., alcoholism, overweight, cancer) and may be also applied to animal populations (e.g., swine influenza), we are concretely interested in **human infectious diseases**.

The use of mathematical tools applied to this science was a great achievement, since it led to the opportunity of making more rigorous predictions, and even reconstructions, of these mentioned dynamics. Nevertheless, before entering the historical motivation and evolution of the mathematical epidemiology, it is necessary to introduce first some concepts that will help us understand better the following pages.

1.1 Basic concepts

Here we present some basic concepts constrained to our context of human infectious diseases.

- **Epidemic:** Rapid spread of a disease to a large number of people in a concrete population within a short period of time.
- **Endemic:** Local persistence of a disease in a geographical area without external inputs.
- **Pandemic:** Spread of an epidemic across a large region, such as multiple continents or worldwide, increasing substantially the number of infected people.
- **Incidence:** Rate of occurrence of a disease in a population within a specified period of time.
- **Prevalence:** Proportion of a population affected by a disease at a specific time.
- **Immunity:** State of a person having adequate biological defenses to fight a disease.
- **Emerging infectious disease (EID):** Infectious disease whose incidence has increased recently (e.g., in few years) and could increase in the near future.
- **Reemerging disease:** Disease that reappears after a significant decline.
- **Latent period:** Time interval between the infection of the individual and when they become capable of infecting other susceptible individuals.
- **Incubation period:** Time interval between the exposure to a disease and the appearance of the first symptoms and clinical signs.
- **Infectious period:** Time interval during which an individual is infectious.

- **Case fatality rate (CFR):** Proportion of people diagnosed with a disease who die from the disease.
- **Infection fatality rate (IFR):** Proportion of people infected by a disease, including asymptomatic and undiagnosed cases, who die from the disease.
- **Basic reproduction number (R_0):** Ratio of secondary infections expected from one infected individual in a completely susceptible population.
- **Effective reproduction number (R_t):** Ratio of secondary infections expected from every infected individual in the current susceptible population.

1.2 Historical motivation

Epidemics have invaded populations since the first recordings of mankind history. Apparently, the most ancient writings about these infectious diseases are in the Bible under the name of *plagues*, and, before Hippocrates (460 B.C. - 370 B.C.) set the first basis on medical science, it was believed that these diseases came from divine anger against the populations [1]. Hippocrates was the first to assert in his *Third Book of the Epidemics* that the spread of these diseases was closely related to the change of season and the humidity in the air. Since then, many epidemics have been devastating populations over the time - [2] between 1347 and 1350, the Black Deaths are estimated to have caused the death of over one third of the European population, and they kept on reemerging during 300 more years; the *Spanish* flu epidemic caused between 1918 and 1919 more than 50 million deaths all over the world; and, finally, even though there have been noticeable improvements on this medical science, the current COVID-19 is estimated to have already killed almost half a million people worldwide in half a year.

John Graunt (1620-1674) was the first person to study the data related to infectious diseases in his book *Natural and Political Observations made upon the Bills of Mortality* (1662), he analyzed the weekly *exitus* records in London and gave a method for comparing risks of dying from various diseases, which set the basis of a theory of competing risks. As these studies were developed, their interest was every time greater due to the nature of these diseases, since they usually suffer from some genomic changes which may make them immune to current vaccines or medicines, and hence it is important to keep on researching and creating new protocols so as to stop their spread as soon as possible. As a curious fact about these protocols, now that most of the world has been quarantined, it is known that the first report of a quarantined place dates from 1377, during the aforementioned Black Plagues, when the Venetian colony Ragusa (the current Dubrovnik in Croatia) imposed a 40-days isolation (which led to the denomination of *quarantine*, due to the Italian word for 40, *quaranta*) to the crews that went to the city for maritime commerce, so as to start avoiding the spread of the disease [3].

Some years after Graunt's studies, it appears that the first mathematical work on epidemiology was thanks to Daniel Bernoulli (1700-1782), during the endemic smallpox disease, who studied if inoculating a mild strain of smallpox in the population would be beneficial in the long term in order to confront the disease and end with its endemic nature. After that, many developments have been made in this area, and we will followingly present several mathematical ways that were set for studying the dimensions of these problems.

1.3 Mathematical modelling in epidemiology

Such as aforementioned, Daniel Bernoulli's study on smallpox is considered the first mathematical study on epidemiology in history [2], but it was not until the 20^{th} century that the basis of the current models was presented.

It was set that diseases spread by contact through some kind of virus or bacterium, and therefore W. H. Hamer proposed in 1906 that this spread should depend on the number of susceptible and infectious individuals, and suggested a *mass-action law* for the rate of new infections, which will be detailed in the next Section with a concrete example. This set the basis of the **compartmental models**, which are the ones we are going to work with. The first known basic compartmental models for our concerning matter in epidemics can be found described in a paper of Kermack and McKendrick from 1927 [4], and now we will present some compartments that are usually taken into account in these models:

- ***M***: Individuals with temporary immunity for the disease; for instance, some children may acquire IgG (immunoglobulin G) antibodies from their mother if she has been infected during her pregnancy. Once lost this immunity, they become susceptible.
- ***S****usceptible*: Healthy individuals who are susceptible of getting infected.
- ***E****xposed*: Individuals in the latent period of the infection with low or null probability of transmitting the infection.
- ***I****nfected*: Individuals already infected by the disease with high probability of transmitting the infection.
- ***R****ecovered*: Individuals who have recovered from the infection.

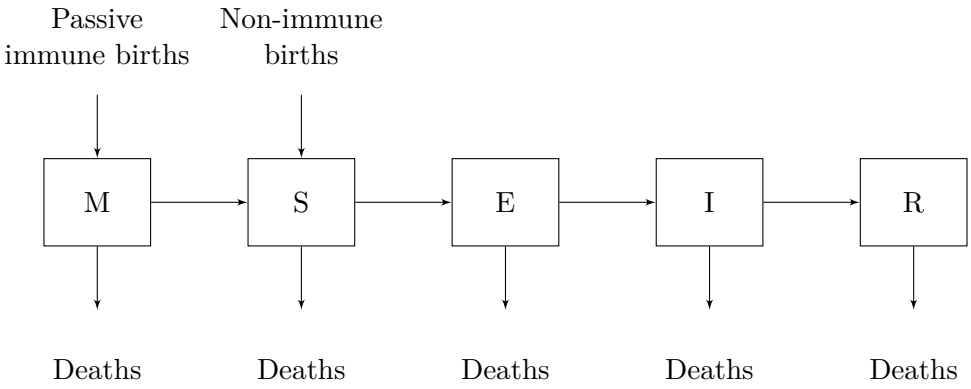


Figure 1: Example of a diagram for a MSEIR model.

These are the basic states, although there may be more such as hospitalized (H), quarantined (Q), vaccinated (V), etc. Given these compartments, there are multiple combinations to build different models depending on the characteristics of the disease, such as the way of transmission

(i.e., by contact, using vectors such as mosquitoes, by air) or the infectious agents (i.e., viruses, bacteria, fungi). In particular, to set up all the basic theory, in Section 2 we will work with a SIRS model, i.e., the infected individuals recover and go back to be susceptible again after some time.

Nevertheless, this is not the only way of studying mathematically an epidemic. In the Kermack-McKendrick models, the authors tend to consider a homogeneous population where there will be some kind of *harmonious* (usually, exponential) behaviour, this is, there is a large enough number of individuals in each compartment so they can be treated homogeneously in terms of age or gender. However, when the number of infected individuals is too low, there is a strong **stochastic component**, since the disease will spread in one way or another depending on the behaviour and characteristics of these individuals. The stochastic models are based on the Galton-Watson process, a result which was first given by Galton and Watson in [5]; when graph theory was developed on the 50's, combined became a strong tool for mathematical epidemiology. Moreover, of course, it can be also considered a **hybrid model** using both techniques.

As many scientific models, these mathematical models may, for instance, help us to:

- Clarify the hypotheses, variables and parameters, and observe the possible changes in the disease dynamics when varying some of them.
- Study, compare and optimize the control measures for the disease.
- Test experimentally some theories and conjectures to have some intuitions on them.
- Identify tendencies, make general predictions or reconstruct the global behaviour of the disease.

Nevertheless, all these models have also some limitations:

- To work mathematically with the epidemic, they must be sufficiently complex to generally adapt to their behaviour, but also simple enough or it will not be possible to treat it theoretical nor numerically. Therefore, they often simplify the reality and do not pretend to model it exactly.
- When we work with deterministic models, our results will be determined and will not consider random events that may interfere with homogeneities.
- On the other hand, considering stochasticity highly increases the computational costs, and hence an equilibrium between both approaches must be found.

In conclusion, the following results must be treated as tools and always resituated in their context. This will help us to clarify the limitations of our models.

2 The SIRS model

To illustrate a typical mathematical model for epidemiology, a study of a basic SIRS model is presented, which will lead us to introduce some new concepts and results that will be useful later for our future study.

First of all, what does the name SIRS stand for? As presented before, each of the letters represent one of the compartments (also called *states*) of the population we are going to work with - in this case, we have the following ones:

- ***S**usceptible*: Healthy individuals who are susceptible of getting infected.
- ***I**nfected*: Individuals infected by the disease.
- ***R**ecovered*: Individuals who have recovered from the infection and are temporarily immune to the disease.

Each of these compartments is a **portion** of the total population. This model is **circular** in the sense that, once recovered, an individual turns to be susceptible again after a period of time.

In this example, although the model will be defined for infinite time, we will effectively work with a time short enough to suppose a normalized constant total population $N = 1$, i.e., $S(t) + I(t) + R(t) = 1$, and will not have into account natural mortality nor any births. Since we have normalized, our compartments S, I and R are dimensionless and represent proportions of the population.

Now, as aforementioned in Section 1.3, to define our system, we will consider a *mass-action law*, which consists on supposing the rate of new infections is directly proportional to the current proportion of infected individuals I/N; this rate, therefore, will be $\beta SI/N$, for some constant β. Because of the previous explanation, this is the same as βSI. Hence, we define the following system of ordinary differential equations:

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t}(t) = -\beta I(t)S(t) + \mu R(t), \\ \dfrac{\mathrm{d}I}{\mathrm{d}t}(t) = \beta I(t)S(t) - \gamma I(t), \\ \dfrac{\mathrm{d}R}{\mathrm{d}t}(t) = \gamma I(t) - \mu R(t), \end{cases} \tag{1}$$

where

- $\beta > 0$ (day^{-1}) is the effective contact rate of the disease,
- $\mu > 0$ (day^{-1}) is the loss of immunity rate,
- $\gamma > 0$ (day^{-1}) is the recovery rate.

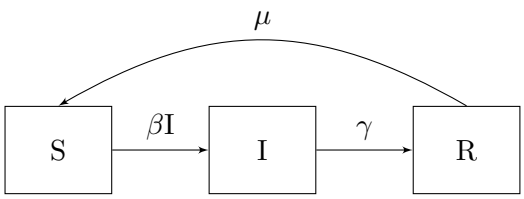


Figure 2: Example of a diagram for a SIRS model.

Recall that the dimensions of our system are consistent due to the previous explanation of the mass-action law, and each equation has dimension [1/T].

Once introduced our model, we can start with the theoretical study. The first step is checking the biological consistency of this model (i.e., the population effectively remains constant and S, I and R are always non-negative). Then, we present more formally the *basic reproduction number* R_0 and a mathematical way for computing it; this will help us to establish the nature of the equilibria of our system, since R_0 will be a threshold number which will determine if the disease becomes endemic or refers. Finally, some numerical results are presented in order to contrast our theory.

2.1 Biological consistency of the model

In this section, we will check some characteristics of our model, observing that our system fulfills our constraint $S(t) + I(t) + R(t) = 1$ (i.e., the population remains constant), for all $t \geq 0$, and is biologically consistent, meaning that we never reach negative proportions of the population; thus, we state:

Theorem 1. *The set*

$$\Omega = \{(S, I, R) \in [0, \infty)^3 : S + I + R = 1\}$$

is ***positively invariant*** *for the system (1), given that* $S(0) + I(0) + R(0) = 1$.

Proof. Let Φ_t be the continuous flux of the system (1). Then, given an initial condition $(S_0, I_0, R_0) \in [0, \infty)^3$, let us see that

$$\Phi_t(S_0, I_0, R_0) \in [0, \infty)^3, \, \forall t \geq 0.$$

To do that, let $(S, I, R) = (X_1, X_2, X_3)$. Then, if $(X_1(\bar{t}), X_2(\bar{t}), X_3(\bar{t})) \in \Omega$, for some $\bar{t} \geq 0$, it is easy to prove that

$$X_i(\bar{t}) = 0 \implies \frac{\mathrm{d}X_i}{\mathrm{d}t}(\bar{t}) \geq 0, \, \forall i = 1, 2, 3.$$

Thus, for any initial condition $(X_1(0), X_2(0), X_3(0)) \in \Omega$, given that our flow Φ_t is continuous, if any of the derivatives is negative, there may exist times t_i, $i = 1, 2, 3$ such that $X_i(t_i) = 0$. Without loss of generality, suppose $t_1 \leq t_2 \leq t_3$; hence,

$$\Phi_{t_1}(X_1(0), X_2(0), X_3(0)) = (0, X_2(t_1), X_3(t_1)) \in \{0\} \times [0, \infty)^2.$$

Now, since $\frac{\mathrm{d}X_1}{\mathrm{d}t}(t_1) \geq 0$ as stated before, X_1 will either increase or remain being 0, and, thus, X_1 remains non-negative.

This is valid for the three of our compartments, so we deduce that Ω is a positive set for our continuous flow $\Phi_{t\geq 0}$.

On the other hand, for any initial solution $(S_0, I_0, R_0) \in \Omega$, $S_0 + I_0 + R_0 = 1$. Besides, adding up the three equations of our system (1), we obtain

$$\frac{\mathrm{d}S}{\mathrm{d}t}(t) + \frac{\mathrm{d}I}{\mathrm{d}t}(t) + \frac{\mathrm{d}R}{\mathrm{d}t}(t) = 0, \, \forall t \geq 0.$$

Hence, $S(t) + I(t) + R(t) = S_0 + I_0 + R_0 = 1$ is a first integral, which leads us to conclude that Ω is positively invariant with respect to the flow. □

2.2 Basic reproduction number

Before starting the mathematical study of the equilibria for (1), we need to present an important threshold number that will help to determine the future of the disease:

Definition 1. *The **basic reproduction number R_0** of a disease is the expected number of cases generated by one infected individual in a completely susceptible population, i.e., the ratio of secondary infections expected from a single case.*

This characteristic number is widely known for establishing an initial criterion to predict the short time evolution of the spread of a disease [6]. Nevertheless, sometimes some part of the population becomes immune to the disease or some control measures are taken to reduce the contact among infected people, and thus this number does not describe the evolution of the disease anymore. Consequently, we define the **effective reproduction number**.

Definition 2. *The **effective reproduction number R_t** of a disease is the expected number of cases generated by one infected individual at a given time t in a partially susceptible population.*

For an autonomous system, we have that $R_t = (S(t)/N)R_0$, and hence computing R_0 already gives us enough information. Nevertheless, for non-autonomous systems it is more complicated and we will see in Section 3 that R_0 will give us less information. However, it is always interesting to compute it and required for most of scientific works on epidemiology.

To compute this number, we are going to use the *Next Generation Matrix* method following the notation of van der Driessche and Watmough [6]. It is named this way since we are aiming to determine the number of secondary infections, i.e., who is the next generation of infected people. Let us describe briefly the method as in the mentioned paper:

We consider a population divided into n compartments (e.g., S, I, R,...). Let $\boldsymbol{p} = (p_1, ..., p_n)^T$ such that $p_i \geq 0$ is the number of individuals in the i^{th} compartment. Suppose there are m compartments with individuals who have the disease active in their body, and $n-m$ who have not got the disease. Then, rearrange and rename these compartments such that $(p_1, ..., p_m) = (x_1, ..., x_m)$ are the m populations with infected individuals and $(p_{m+1}, ..., p_n) = (y_1, ..., y_{n-m})$ are the $n - m$ disease-free compartments. Let $\boldsymbol{P}_f$ be the set of populations in disease-free state

$$\boldsymbol{P}_f = \{\boldsymbol{p} = (\boldsymbol{x}, \boldsymbol{y}) \geq \boldsymbol{0} \mid x_i = 0, i = 1, ..., m\}.$$

Next, let us define the following functions under the assumption that they will be $\mathcal{C}^2$:

- $F_i(\boldsymbol{p})$: rate of appearance of new infections in $i = 1, ..., m$.
- $V_i^+(\boldsymbol{p})$: rate of transfer of individuals to i by other means than infection (for instance, for transition rate).
- $V_i^-(\boldsymbol{p})$: rate of transfer of individuals out of i.

Given this, we can rewrite our epidemiological model as follows:

$$\begin{cases} \dot{p}_i = f_i(\boldsymbol{p}) = F_i(\boldsymbol{p}) - V_i(\boldsymbol{p}),\ i = 1, ..., n, \\ \text{Non-negative initial conditions,} \end{cases} \qquad V_i = V_i^- - V_i^+. \tag{2}$$

We impose the described functions also fulfill the following assumptions:

1. Since our functions F_i, V_i^+ and V_i^- are positive transfers of individuals, each function will be non-negative. Thus,

$$\boldsymbol{p} \geq 0 \implies F_i, V_i^-, V_i^+ \geq 0,\ \forall i = 1, ..., n.$$

2. An empty compartment cannot transfer any individuals out of it, hence

$$p_i = 0 \implies V_i^- = 0.$$

Particularly,

$$\boldsymbol{p} \in \boldsymbol{P}_f \implies V_i^- = 0,\ \forall i = 1, ..., m.$$

Note: Satisfying conditions **1.** and **2.** for (2) leads to the existence of a unique non-negative solution for each non-negative initial condition (see [7]).

3. Since the incidence of infection for uninfected compartments is zero, $F_i = 0,\ \forall i > m$.

4. To maintain $\boldsymbol{P}_f$ invariant, we assume that a population free of disease will remain free of disease. Thus,

$$\boldsymbol{p} \in \boldsymbol{P}_f \implies F_i(\boldsymbol{p}) = V_i^+(\boldsymbol{p}) = 0,\ \forall i = 1, ..., m.$$

5. We want an equilibrium point $\boldsymbol{p}_f$ of the disease free state set $\boldsymbol{P}_f$ to be (locally asymptotically) stable, i.e., if we remain close to $\boldsymbol{p}_f$, the population will tend to the disease free equilibrium (DFE). Taking into account the linearized system

$$\dot{\boldsymbol{p}} = Df(\boldsymbol{p}_f)(\boldsymbol{p} - \boldsymbol{p}_f), \tag{3}$$

we ask that

$$F(\boldsymbol{p}) = \boldsymbol{0} \implies eig(Df(\boldsymbol{p}_f)) \subset \{(a + bi) \in \mathbb{C} \,|\, a < 0\},$$

this is, we want $\boldsymbol{p}_f$ to be stable in absence of new infection. Be aware that $\boldsymbol{p}_f$ is usually on the boundary of our domain, and hence we should work with one sided derivatives.

Given these assumptions, we state the following result from [6]:

Lemma 1. *If $\boldsymbol{p}_f$ is a DFE of (2) and $f_i(\boldsymbol{p})$ satisfies the aforementioned assumptions, then*

$$DF(\boldsymbol{p}_f) = \begin{pmatrix} \mathcal{F} & 0 \\ 0 & 0 \end{pmatrix}, \quad DV(\boldsymbol{p}_f) = \begin{pmatrix} \mathcal{V} & 0 \\ J_3 & J_4 \end{pmatrix},$$

where

- *$\mathcal{F}$ and $\mathcal{V}$ are $m \times m$ matrices such that*

$$\mathcal{F} = \left[\frac{\partial F_i}{\partial p_j}(\boldsymbol{p}_f)\right]_{\substack{i=1\ldots,m,\\ j=1,\ldots,m}} \text{ is non-negative}, \quad \mathcal{V} = \left[\frac{\partial V_i}{\partial p_j}(\boldsymbol{p}_f)\right]_{\substack{i=1\ldots,m,\\ j=1,\ldots,m}} \in GL(m,\mathbb{R}),$$

- *$\mathcal{V}$ is an M-matrix, i.e., its non-diagonal entries are non-positive and its eigenvalues have non-negative real part (positive real part, indeed, due to the invertibility),*
- *J_3 and J_4 are an $(n-m) \times m$ and an $(n-m) \times (n-m)$ matrices, respectively, such that*

$$eig(J_4) \subset \{(a+ib) \in \mathbb{C} \,|\, a > 0\}.$$

Now, to determine the infections produced by one individual, let us consider again the linearized system with no reinfections introduced in **5.**

$$\dot{\boldsymbol{p}} = -DV(\boldsymbol{p}_f)(\boldsymbol{p} - \boldsymbol{p}_f).$$

Recall that we asserted in this case $\boldsymbol{p}_f$ is locally asymptotically stable. From here, we are going to determine the local behaviour when introducing a small number of infected individuals.

Let $r_i(0)$ be the initial number of infected individuals in compartment i and denote $\boldsymbol{r}(t) = (r_i(t))_{i=1,\ldots,m}$ the number of these initially infected individuals in the infected compartments after t time units, i.e., $\boldsymbol{r}(t)$ represents the first m components of $\boldsymbol{p}$. Recalling the form of $DV(\boldsymbol{p}_f)$ in the previous Lemma, we have the equation for $\boldsymbol{r}(t)$

$$\dot{\boldsymbol{r}}(t) = -\mathcal{V}\boldsymbol{r}(t),$$

whose solution is $\boldsymbol{r}(t) = e^{-\mathcal{V}t}\boldsymbol{r}(0)$.

Attending to Lemma 1, $\mathcal{V}$ is invertible with all its eigenvalues having positive real part; hence, considering the linearized system (3) and integrating $\mathcal{F}\boldsymbol{r}(t)$ from zero to infinite, it is obtained the expected number of new infections due to the initially infected individuals as $\mathcal{F}\mathcal{V}^{-1}\boldsymbol{r}(0)$. The fact that $\mathcal{V}$ is a non-singular M-matrix is equivalent to the fact that it is positive-inverse, i.e., $\mathcal{V}^{-1}$ has all positive entries. This, added up to Lemma 1, leads to $\mathcal{F}\mathcal{V}^{-1}$ positive. Let us now interpret what each entry of this last matrix means:

Suppose an infected individual entering to compartment k. $\mathcal{V}^{-1}(j,k)$ corresponds to the average period of time this individual will spend in compartment j before leaving. On the other hand, $\mathcal{F}(i,j)$ corresponds to the transmission rate from compartment j to i. Thus, $\mathcal{F}\mathcal{V}^{-1}(i,k)$ corresponds to

the **expected number of infections in i from the infected individual introduced in k**. Van der Driessche and Watmough based themselves on the first study of this method by Diekmann, Heesterbeek and Metz in 1990 (see [8]) to conclude that $\mathcal{F}\mathcal{V}^{-1}$ is the so-called Next Generation Matrix and R_0 is defined as

$$R_0 = \rho(\mathcal{F}\mathcal{V}^{-1}), \tag{4}$$

being $\rho(\mathcal{F}\mathcal{V}^{-1})$ the spectral radius of the matrix, which makes sense since we are supposing that the persistence of the disease will depend on the largest of these expected numbers.

Let us now compute R_0 for our particular model (1). Since $m = 1$, as the only infected population is I, we rewrite our equations as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} I(t) \\ S(t) \\ R(t) \end{pmatrix} = F(I,S,R) - V(I,S,R), \tag{5}$$

being

$$F(I,S,R) = \begin{pmatrix} \beta IS \\ 0 \\ 0 \end{pmatrix} \text{ and } \quad V(I,S,R) = \begin{pmatrix} \gamma I \\ \beta IS - \mu R \\ -\gamma I + \mu R \end{pmatrix}.$$

Differentiating both functions with respect to (I,S,R), we obtain

$$DF(I,S,R) = \begin{pmatrix} \beta S & \beta I & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } \quad DV(I,S,R) = \begin{pmatrix} \gamma & 0 & 0 \\ \beta S & \beta I & -\mu \\ -\gamma & 0 & \mu \end{pmatrix}.$$

Next, we determine the disease free equilibrium (DFE) of our model, i.e., we look for a point $\boldsymbol{p}_f = (0, S_f, R_f)$ such that $\frac{\mathrm{d}I}{\mathrm{d}t} = \frac{\mathrm{d}S}{\mathrm{d}t} = \frac{\mathrm{d}R}{\mathrm{d}t} = 0$. Thus, substituting in (5) by a point satisfying this criterion, we have that

$$\begin{cases} 0 = 0, \\ 0 = \mu R_f, \\ 0 = -\mu R_f, \end{cases}$$

and we obtain $(0, S_f, 0)$. Attending to our constraint $S + I + R = 1$, we get

$$\boldsymbol{p}_f = (0,1,0).$$

Considering this point in the Jacobian matrices, we have that

$$DF(\boldsymbol{p}_f) = \begin{pmatrix} \beta & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } \quad DV(\boldsymbol{p}_f) = \begin{pmatrix} \gamma & 0 & 0 \\ \beta & 0 & -\mu \\ -\gamma & 0 & \mu \end{pmatrix},$$

from which we deduce, taking into account that $m = 1$, that the basic reproduction number is given by

$$R_0 = \rho(\mathcal{F}\mathcal{V}^{-1}) = \beta/\gamma.$$

Remark: One can observe that the condition of Lemma 1 for the eigenvalues of J_4 is not fulfilled, which is due to the fact that our three equations are linearly dependent due to the aforementioned constraint. Later, we will reduce this system in order to make our computations easier.

On the other hand, we can find another equilibrium point $\boldsymbol{p}_e = (I_e, S_e, R_e)$ for our system such that $I > 0$. This point will be our **endemic equilibrium**, i.e., the equilibrium for which the disease persists.

$$\begin{cases} 0 = \beta S_e - \gamma, \\ 0 = -\beta I_e S_e + \mu R_e = -\beta I_e S_e + \mu(1 - S_e - I_e), \\ 0 = \gamma I_e - \mu R_e = \gamma I_e - \mu(1 - S_e - I_e), \end{cases} \implies S_e = \frac{\gamma}{\beta},\ I_e = \frac{\mu(1-\gamma/\beta)}{\gamma+\mu},\ R_e = \frac{\gamma(1-\gamma/\beta)}{\gamma+\mu},$$

which, rewritten in terms of R_0, leads us to the endemic equilibrium

$$\boldsymbol{p}_e = \left(\frac{\mu(1-1/R_0)}{\gamma+\mu}, \frac{1}{R_0}, \frac{\gamma(1-1/R_0)}{\gamma+\mu}\right) = \left(I_e, S_e, \frac{\gamma}{\mu} I_e\right).$$

Notice that this equilibrium point is only admissible (and different from $\boldsymbol{p}_f$) for $R_0 > 1$.

2.3 Stability of the equilibria

The next step is studying the stability of these two points, $\boldsymbol{p}_f$ and $\boldsymbol{p}_e$, with respect to the value of R_0. To do that, let us recall first some definitions regarding Lyapunov's stability theory.

Let us consider the following autonomous non-linear system:

$$\begin{cases} \dot{\boldsymbol{p}} &= f(\boldsymbol{p}(t)), \\ \boldsymbol{p}(0) &= \boldsymbol{p}_0, \end{cases} \tag{6}$$

where $\boldsymbol{x}(t) \in D \subseteq \mathbb{R}^n$, D open, and $f : D \to \mathbb{R}^n$ a continuous function. Suppose the system has an equilibrium point $\boldsymbol{p}_s$.

Definition 3. *The equilibrium point $\boldsymbol{p}_s$ is said to be **Lyapunov stable** if, for all $\varepsilon > 0$, $\exists\, \delta > 0$ such that $\|\boldsymbol{p}(0) - \boldsymbol{p}_s\| < \delta \implies \|\boldsymbol{p}(t) - \boldsymbol{p}_s\| < \varepsilon$, for all $t > 0$.*

Definition 4. *The equilibrium point $\boldsymbol{p}_s$ is said to be **asymptotically stable** if there exists $\delta > 0$ such that $\|\boldsymbol{p}(0) - \boldsymbol{p}_s\| < \delta \implies \lim_{t\to\infty} \|\boldsymbol{p}(t) - \boldsymbol{p}_s\| = 0$.*

Definition 5. *The equilibrium point $\boldsymbol{p}_s$ is said to be **unstable** if it is not stable, i.e., if there exists an $\varepsilon > 0$ such that, $\forall \delta > 0$ and $\|\boldsymbol{p}(0) - \boldsymbol{p}_s\| < \delta$, $\exists t_0 > 0 \,|\, \|\boldsymbol{p}(t) - \boldsymbol{p}_s\| > \varepsilon$, $\forall t > t_0$.*

Given this, we state the following assertion:

Theorem 2. *The two equilibria of system (2) fulfill the next stability behaviours:*

1. *$\boldsymbol{p}_f = (0, 1, 0)$ is **globally asymptotically stable (GAS)** if $R_0 \leq 1$ and **unstable** if $R_0 > 1$.*
2. *$\boldsymbol{p}_e = (I_e, S_e, \frac{\gamma}{\mu} I_e)$, being*

$$I_e = \mu \frac{(1 - 1/R_0)}{\gamma + \mu}, \quad S_e = \frac{1}{R_0},$$

*is **locally asymptotically stable (LAS)** for $R_0 > 1$ (and not feasible for $R_0 \leq 1$).*

To prove this, we will first present some theoretical results that will help us. The part for the global stability will be based on the study proposed by van der Driessche and Shuai in [9], including the use of LaSalle's Invariance Principle [10]. On the other hand, we will use Hartman-Grobman's theorem for the local results [11]. Thus, let us state this theory.

Definition 6 (Sharp threshold property)**.** *Suppose a model of the form* (2) *and* R_0 *computed by the NGM, this model has the* ***sharp threshold property*** *if* R_0 *is such that*

- $\boldsymbol{p}_f$ *is GAS for* $R_0 \leq 1$*, and*
- $\exists! \, \boldsymbol{p}_e$ *GAS in the interior of the feasible region for* $R_0 > 1$*.*

Indeed, this is a more complete result than the one we want to prove, already knowing that this endemic equilibrium exists. Nevertheless, stating an equilibrium is globally asymptotically stable may be a tough labor for complex models. For biological models, Lyapunov functions are very useful, although they may not be easy to build.

Definition 7 (**Lyapunov function**)**.** *Suppose* $\boldsymbol{p}_0$ *is an equilibrium for system (6). Then, a scalar function* $V : \mathbb{R}^n \to \mathbb{R}$ *is a* ***Lyapunov function*** *for this equilibrium if* V *is continuous,* $V(\boldsymbol{p}_0) = 0$*, it is locally positive-definite (i.e., there exists a* $\boldsymbol{p}_0$*-neighbourhood* $D \subset \mathbb{R}^n$ *such that* $V(\boldsymbol{p}) > 0$*, for all* $\boldsymbol{p} \in D \setminus \{\boldsymbol{p}_0\}$*) and* $\dot{V}(\boldsymbol{p}) = \nabla V \cdot f(\boldsymbol{p}) \leq 0$*, for all* $\boldsymbol{p} \in D$*.*

It is clear from the definition that $\dot{V}(\boldsymbol{p}_0) = 0$. If we are able to find a Lyapunov function, we have the following theorem to state the stability of the equilibrium.

Theorem (**Lyapunov stability** [12])**.** *Let* $\boldsymbol{p}_0$ *be an equilibrium for system (6), and let* V *be a Lyapunov function for this point. Then:*

- $\boldsymbol{p}_0$ *is* ***stable*** *for system (6).*
- *If there exists an* $\boldsymbol{p}_0$*-neighbourhood* $B \subset D$ *such that* $\dot{V}(\boldsymbol{p}) < 0$*, for all* $\boldsymbol{p} \in B \setminus \{\boldsymbol{p}_0\}$*, then* $\boldsymbol{p}_0$ *is* ***locally asymptotically stable****.*
- *If* V *is* ***globally*** *positive-definite (i.e.,* $V(\boldsymbol{p}) > 0$*, for all* $\boldsymbol{p} \in \mathbb{R}^n \setminus \{\boldsymbol{p}_0\}$*) and* $\dot{V}(\boldsymbol{p}) < 0$*, for all* $\boldsymbol{p} \in \mathbb{R}^n \setminus \{\boldsymbol{p}_0\}$*, then* $\boldsymbol{p}_0$ *is* ***globally asymptotically stable****.*

To prove the global asymptotical stability for $\boldsymbol{p}_f$, van der Driessche and Shuai present a matrix-theoretic method based on the Perron-Frobenius theorem, which asserts that a real square positive matrix has a unique largest real eigenvalue for which its corresponding eigenvector may be chosen with all its entries positive. On the other hand, for $\boldsymbol{p}_e$, they present a graph-theoretic method which, unfortunately, in our case we have not been able to make it work, and thus we will not introduce this second theory.

$\boldsymbol{p}_f$ GAS, matrix-theoretic method:

To perform this method, we will follow a guide to construct a *Lyapunov-type* function, i.e., a locally positive-definite and non-increasing function. It is necessary to clarify that it will not be a

proper Lyapunov function since we are working in a compact set, and therefore the original theorem would not be directly applicable. Indeed, it is usual working in compact sets in mathematical epidemiology given that we are working with populations that cannot be negative quantities and are sometimes constrained by a constant number of individuals. Hence, recalling $\boldsymbol{p} = (\boldsymbol{x}, \boldsymbol{y})$, set first

$$f(\boldsymbol{x}, \boldsymbol{y}) := (\mathcal{F} - \mathcal{V})\boldsymbol{x} - F(\boldsymbol{x}, \boldsymbol{y}) + V(\boldsymbol{x}, \boldsymbol{y}).$$

Notice that we are making an abuse of notation and now we are restricting F and V to the infected compartments (i.e., $\boldsymbol{x}$). For these compartments, we have

$$\dot{\boldsymbol{x}} = (\mathcal{F} - \mathcal{V})\boldsymbol{x} - f(\boldsymbol{x}, \boldsymbol{y}). \tag{7}$$

From previous assumptions, $f(\mathbf{0}, \boldsymbol{y}) = \mathbf{0}$. Let $\nu^T \geq 0$ be the left eigenvector of $\mathcal{V}^{-1}\mathcal{F}$, corresponding to the eigenvalue R_0. Then, we present a general method to construct a Lyapunov function:

Theorem 3 ([9]). *Let $\mathcal{F}, \mathcal{V}$ and $f(\boldsymbol{x}, \boldsymbol{y})$ as previously defined. If $f(\boldsymbol{x}, \boldsymbol{y}) \geq \mathbf{0}$ in $\Omega \subset \mathbb{R}^n_{0+}$, $\mathcal{F} \geq \mathbf{0}$, $\mathcal{V}^{-1} \geq \mathbf{0}$ and $R_0 \leq 1$, then $Q(\boldsymbol{x}) = \nu^T \mathcal{V}^{-1} \boldsymbol{x}$ is a Lyapunov-type function for our model (7).*

Proof. It is clearly locally positive-definite for $\boldsymbol{x} \neq \mathbf{0}$. If we now differentiate:

$$\dot{Q}(\boldsymbol{x}) = \nu^T \mathcal{V}^{-1} \dot{\boldsymbol{x}} = \nu^T \mathcal{V}^{-1} \left((\mathcal{F} - \mathcal{V})\boldsymbol{x} - f(\boldsymbol{x}, \boldsymbol{y}) \right) = \nu^T (R_0 - 1)\boldsymbol{x} - \nu^T \mathcal{V}^{-1} f(\boldsymbol{x}, \boldsymbol{y}).$$

The last term is non-positive, since $\nu^T, \mathcal{V}^{-1} \geq \mathbf{0}$, $f(\boldsymbol{x}, \boldsymbol{y}) \geq \mathbf{0}$ in Ω. Besides, given that $R_0 \leq 1$, this implies $\dot{Q}(\boldsymbol{x}) \leq 0$ in Ω. Hence, Q is Lyapunov-type for our system. □

Theorem (**LaSalle's Invariance principle**, [10]). *Let V be a Lyapunov-type function of the system in a compact positively invariant set Ω and let $\boldsymbol{x}(t) = \phi(t, \boldsymbol{x}_0)$ be a solution that remains in Ω for $0 \leq t < t^*$, for some $t^* > 0$. Suppose there exists a constant $c \in \mathbb{R}^+$ such that $V(\boldsymbol{x}) \leq c$ in Ω and let $\Gamma = \{\boldsymbol{x} \in \Omega \,|\, \dot{V}(\boldsymbol{x}) = 0\}$ and $\Gamma_0 \subset \Gamma$ its largest invariant subset. Then, any solution $\boldsymbol{x}(t)$ that is precompact (i.e., its closure is compact), tends to Γ_0 when $t \to \infty$.*

Theorem (**Hartman-Grobman**, [11]). *Suppose system (6) and let $\boldsymbol{x}_s$ be a hyperbolic equilibrium point (i.e., the Jacobian of the system evaluated at this point has no eigenvalue with real part equal to zero). Then, the local flow around $\boldsymbol{x}_s$ is topologically conjugate to the linear flow $\dot{\boldsymbol{x}} = J(\boldsymbol{x}_s)(\boldsymbol{x} - \boldsymbol{x}_s)$ around $\boldsymbol{x}_s$, being J the Jacobian of the system.*

Once established all these tools, we prove **Theorem 2**.

Proof. • $\boldsymbol{p}_f$:

Let us start rewriting our system as in (7), i.e.,

$$\dot{x} = (\mathcal{F} - \mathcal{V})x - f(x, \boldsymbol{y}),$$

where

$$f(x, \boldsymbol{y}) = \beta I(1 - S).$$

Since $(I,S,R) \in [0,1]^3$, $f(I,(S,R)) \geq 0$ in Ω. Now, let $\nu^T \geq 0$ be the left eigenvector of $\mathcal{V}^{-1}\mathcal{F}$ corresponding to the eigenvalue R_0, i.e., a row vector such that $\nu^T\mathcal{V}^{-1}\mathcal{F} = R_0\nu^T$. Let us notice that, if R_0 is an eigenvalue for $\mathcal{F}\mathcal{V}^{-1}$, so it is for $\mathcal{V}^{-1}\mathcal{F}$, since

$$\det(\mathcal{F}\mathcal{V}^{-1} - R_0 I_m) = 0 \implies \det(\mathcal{F} - R_0\mathcal{V}) = 0 \implies \det(\mathcal{V}^{-1}\mathcal{F} - R_0 I_m) = 0.$$

Therefore, given that $\nu^T\mathcal{V}^{-1}\mathcal{F} = \beta/\gamma = R_0$, let us choose $\nu^T = 1$. Thus, attending to Theorem 3,

$$Q(x) = \nu^T\mathcal{V}^{-1}I = \frac{I}{\gamma}$$

is a Lyapunov-type function for our system.

To prove now the global asymptotic stability, we need to apply LaSalle's invariance principle:

We have Q, a Lyapunov-type function in the compact set Ω, and we know that any solution $x(t)$ remains in Ω, i.e., it is bounded (which is equivalent to being precompact in this case). Besides, $Q(x) \leq 1/\gamma$ in Ω. Let us see who Γ is in our case, taking into account that $R_0 \leq 1$:

$$\dot{Q}(x) = 0 \implies f(x,\boldsymbol{y}) = \beta I(1-S) = 0,$$

i.e., $I = 0$ and $S + R = 1$ or $S = 1$ and $I = R = 0$. Hence, the maximal invariant subset is $\Gamma_0 = \{(0,S,R) \in \Omega\}$. If we now suppose $R > 0$, we obtain the equation for S

$$\frac{\mathrm{d}S}{\mathrm{d}t}(t) = \mu(1 - S(t)),$$

which has solution $S(t) = 1 + S_0 e^{-\mu t}$, but then $S(t) > 1$, which is in contradiction with the fact that Ω is invariant with respect to the flow. Thus, the remaining possibility is $\Gamma_0 = \{\boldsymbol{p}_f = (0,1,0)\}$. from where we conclude that any solution $x(t)$ will tend to x_f when $t \to \infty$, which means that x_f is GAS for $R_0 \leq 1$, and so it is $\boldsymbol{p}_f$.

On the other hand, the proof of the global instability of $\boldsymbol{p}_f$ for $R_0 > 1$ can be found in Theorem 2 of [6]. However, we can easily prove the local instability, which for this case is already an important result, because that would mean that the disease would become persistent, independently on the global behaviour of the solution. To do this, we are going to use Hartman-Grobman's theorem. But, first, notice that, if we work directly with our original system, we will always obtain a zero eigenvalue, since there is a constraint relating them. Thus, let us rewrite the system taking into account that $S(t) + I(t) + R(t) = 1$, for all $t > 0$:

$$\begin{cases} \dfrac{\mathrm{d}I}{\mathrm{d}t} = \beta SI - \gamma I, \\ \dfrac{\mathrm{d}S}{\mathrm{d}t} = -\beta SI + \mu R, \\ \dfrac{\mathrm{d}R}{\mathrm{d}t} = \gamma I - \mu R, \\ S + I + R = 1, \end{cases} \implies \begin{cases} \dfrac{\mathrm{d}I}{\mathrm{d}t} = \beta SI - \gamma I, \\ \dfrac{\mathrm{d}S}{\mathrm{d}t} = -\beta SI + \mu(1 - S - I), \end{cases} \tag{8}$$

for which we have the restricted disease-free equilibrium $\overline{\boldsymbol{p}}_f = (0,1)$. Now, we will compute the Jacobian of this system on $\overline{\boldsymbol{p}}_f$ in order to see if Hartman-Grobman may be applied:

$$J(\overline{\boldsymbol{p}}_f) = \begin{pmatrix} \beta - \gamma & 0 \\ -\beta - \mu & -\mu \end{pmatrix}.$$

Given that $R_0 = \beta/\gamma > 1$ by hypothesis, we obtain $\beta - \gamma > 0$. Hence, we obtain both one negative and one positive eigenvalues, i.e., $\overline{\boldsymbol{p}}_f$ is locally a saddle-point, which means that there is still a stable manifold such that every solution starting (locally) on this manifold will tend to our disease-free equilibrium. This stable manifold must be tangent to the linear space $E^s = \langle (0,1) \rangle$, since $(0,1)$ is the eigenvector associated to $-\mu$. Moreover, our stable manifold **is** E^s, since it must be a 1D curve and any solution with initial condition $(0, S^*)$, $S^* \in (0,1)$, tends to $\overline{\boldsymbol{p}}_f$, given that $\mathrm{d}I/\,\mathrm{d}t = 0$, $\mathrm{d}S/\,\mathrm{d}t > 0$, for all $t > 0$. Concretely, $\mathrm{d}S/\,\mathrm{d}t$ would be less or equal than zero in this case only for $S \geq 1$, but, in particular, we would have already reached the equilibrium $(I = 0, S = 1)$.

In conclusion, any solution with initial condition (I^*, S^*), $I^* > 0$, close to $\overline{\boldsymbol{p}}_f$ will move away when time goes to infinite and, as previously said, the disease will not refer.

- $\boldsymbol{p}_e$:

We have already seen that $\boldsymbol{p}_e$ is only feasible for $R_0 > 1$, so let us see what happens for this case.

In [9], the authors present a graph-theoretic method, but we have not been able to find appropriate functions for this theory to work in this case. Hence, we are going to prove local asymptotic stability using Hartman-Grobman again. For that, we use again our restricted system (8) and compute its Jacobian on $\overline{\boldsymbol{p}}_e = (\mu(1-1/R_0)/(\gamma+\mu), 1/R_0)$ (recall that we are in the case $R_0 = \beta/\gamma > 1$):

$$J(\boldsymbol{p}_e) = \begin{pmatrix} \beta\dfrac{1}{R_0} - \gamma & \beta\mu\dfrac{1-1/R_0}{\gamma+\mu} \\ \\ -\beta\dfrac{1}{R_0} - \mu & -\beta\mu\dfrac{1-1/R_0}{\gamma+\mu} - \mu \end{pmatrix} = \begin{pmatrix} 0 & \mu\dfrac{\beta-\gamma}{\gamma+\mu} \\ \\ -\gamma - \mu & -\mu\dfrac{\beta+\mu}{\gamma+\mu} \end{pmatrix}.$$

Let us compute its eigenvalues in order to see if we can apply Hartman-Grobman's theorem:

$$\begin{vmatrix} -\lambda & \mu\dfrac{\beta-\gamma}{\gamma+\mu} \\ \\ -\gamma-\mu & -\mu\dfrac{\beta+\mu}{\gamma+\mu} - \lambda \end{vmatrix} = \lambda\left(\mu\frac{\beta+\mu}{\gamma+\mu} + \lambda\right) + \mu(\beta-\gamma) = 0 \implies$$

$$\implies \lambda = \frac{-\mu(\beta+\mu)/(\gamma+\mu) \pm \sqrt{\mu^2(\beta+\mu)^2/(\gamma+\mu)^2 - 4(\beta-\gamma)}}{2}.$$

We can study the possible eigenvalues for two different cases for the term under the square root:

- $\mu^2(\beta+\mu)^2/(\gamma+\mu)^2 - 4(\beta-\gamma) \geq 0$: In this case, the square root is real. Besides, since $R_0 = \beta/\gamma > 0$, we have $4(\beta-\gamma) > 0$ and, thus,

$$\sqrt{\mu^2(\beta+\mu)^2/(\gamma+\mu)^2 - 4(\beta-\gamma)} < \sqrt{\mu^2(\beta+\mu)^2/(\gamma+\mu)^2} = \mu^2(\beta+\mu)^2/(\gamma+\mu)^2,$$

  which implies that both eigenvalues would be negative. Hence, we are at a locally asymptotically stable (LAS) point.

- $\mu^2(\beta+\mu)^2/(\gamma+\mu)^2 - 4(\beta-\gamma) < 0$: This implies that the square root is imaginary and, again, the real part of both of our eigenvalues would be negative, leading to a LAS point.

Combining this result with the one for $\overline{\boldsymbol{p}}_f$, $R_0 > 1$, we obtain that the the number of infected people will never be zero and will often tend to an equilibrium.

□

2.4 Numerical study

In this section, we show two examples that illustrate our theoretical results. More precisely, we suppose two different diseases, Disease 1 and Disease 2, with different values of β, and compare their evolution with respect to the exposed results.

Since our $3D$ system has a constraint, $S+I+R=1$, we can plot our results in the Susceptible-Infected plane, such that $S+I \leq 1$.

- **Disease 1**: In Figure 3, we plot the trajectories of a disease associated to a SIRS model with parameters $(\beta,\gamma,\mu) = (0.05, 0.2, 0.2)$. These parameters lead to the reproduction number $R_0 = 0.05/0.2 = 0.25 < 1$, i.e., the disease refers. Indeed, consistently with the exposed theory, our solution trajectories tend to the disease-free equilibrium $(S=1, I=0)$.

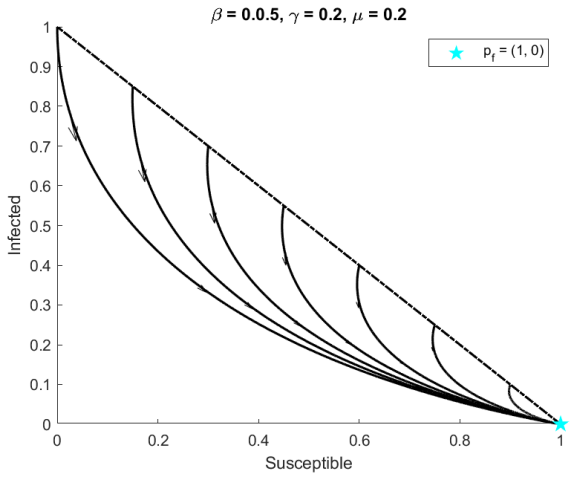


Figure 3: Numerical result for the SIRS model with parameters $(\beta,\gamma,\mu) = (0.05, 0.2, 0.2)$.

- **Disease 2**: In Figure 4, we plot the trajectories of a disease associated to a SIRS model with parameters $(\beta, \gamma, \mu) = (0.3, 0.2, 0.2)$. These parameters lead to the basic reproduction number $R_0 = 0.3/0.2 = 1.5 > 1$, i.e., the disease should tend to the endemic equilibrium.

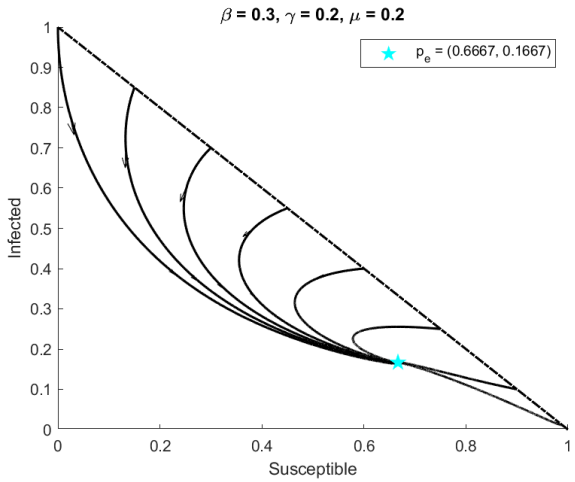


Figure 4: Numerical result for the SIRS model with parameters $(\beta, \gamma, \mu) = (0.3, 0.2, 0.2)$.

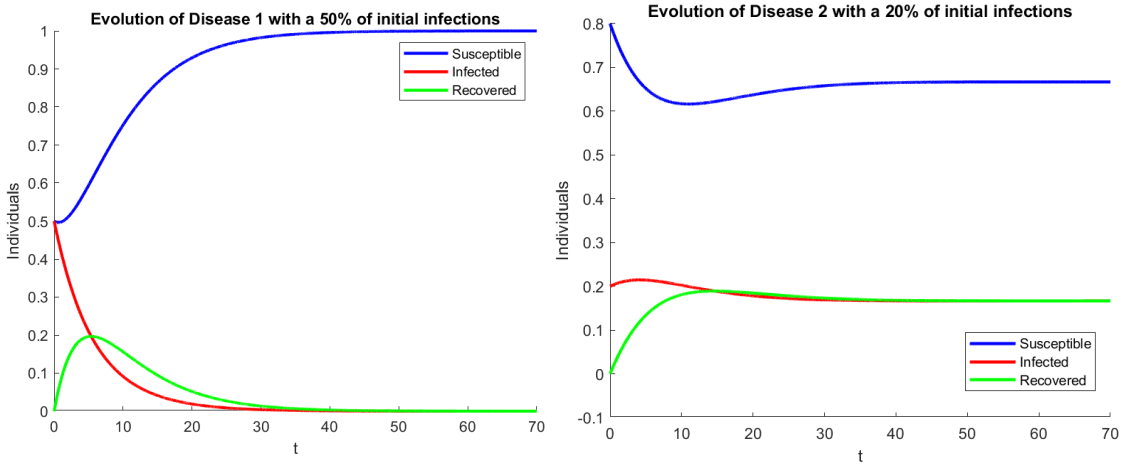


Figure 5: One particular case for each disease. On the left, evolution of Disease 1 with an initial amount of 50% of infections; on the right, evolution of Disease 2 with an initial amount of 20% of infections.

Remember that the endemic equilibrium was on the form

$$\boldsymbol{p}_e = (S_e, I_e) = \left(\frac{1}{R_0}, \frac{\mu(1 - 1/R_0)}{\gamma + \mu} \right),$$

which, in our case, should correspond to

$$\boldsymbol{p}_e \approx (0.6667, 0.1667).$$

In fact, our results match with our theory, since we can observe that our trajectories approach to this point $\boldsymbol{p}_e$.

We note the importance of this threshold number regarding each previous example. In Figure 5, we present the evolution of each compartment, S, I and R, for each set of parameters, given an initial condition. For Disease 1 (which has $R_0 < 1$), although we start with a 50% of the population infected, it rapidly refers. On the other hand, when starting with a 20% of infections in Disease 2 ($R_0 > 1$), we tend to our theoretical endemic equilibrium (notice that, for this case, the limits for R and I coincide since $R_e = (\gamma/\mu)I_e$ and $\gamma = \mu = 0.2$).

3 Epidemiological model for COVID-19

In this section, we study a compartmental θ-SEIHRD model proposed for describing the novel coronavirus disease, denoted by COVID-19, caused by the virus SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2). To construct this model, we will consider the studies performed by B. Ivorra, A. M. Ramos et al. in [13]. We do a theoretical study analogous to the one presented in the previous section and then adapt the methods proposed in the mentioned paper to reconstruct the evolution of the disease in the Metropolitan Area of Barcelona (AMB, in its name in Catalan).

3.1 Formulation of the model

For this model, we are going to take into account initially 9 different compartments:

- Susceptible (S): Healthy individuals susceptible of being infected.
- Exposed (E): Infected individuals during their incubation period, they do not present any symptoms and can infect other people, but with lower probability that the ones in the infectious compartments. Once finished the incubation period, they pass to one of the infectious compartments. In the literature, they are also called *presymptomatics*.
- Infectious (I): Infected individuals who present symptoms and may infect other people.
- Undetected infectious (I_u): Infected individuals who are asymptomatic or present mild or medium symptom and may infect other people but will not be diagnosed. After this period, they will directly recover.
- Hospitalized or quarantined at home (H_R): Detected infected individuals that may still infect other people, they will recover.
- Severe hospitalization (H_D): Hospitalized individuals that may still infect other people and will die.
- Detected recovered (R_d): Detected individuals who have recovered from the infection, cannot infect anymore and have developed immunity.
- Undetected recovered (R_u): Undetected individuals who have recovered from the infection, cannot infect anymore and have developed immunity.
- Dead (D): Individuals who have not survived the infection, not infectious anymore.

There may be other compartments, depending on the imposition of the authorities or the sanitary developments (for example, in a near future we could consider a Vaccinated compartment), but we will not study these cases for the lack of information.

Our model will focus on the evolution of the disease in the AMB, Barcelona, Spain. The first reported case was on February 27^{th}, and on March 11^{th} the World Health Organization declared the disease to be a **pandemic** [14]. This declaration led the Spanish government to approve the **alarm**

state on March 14^{th} to be effective on March 15^{th} at 12 a.m. [15] and put in quarantine the whole country. On March 28^{th}, this quarantine was hardened and the workers of no-essential services were also obligated to stay home. The first step of the de-escalation was on April 12^{th} when some workers of no-essential services were allowed to go back to work. Two weeks later, on April 26^{th}, kids were allowed to go out with their parents, and then on May 2^{nd} people were allowed exercise outside, always on their own and during some determined hours [16]. On May 13^{th}, the *Instituto de Salud Carlos III* presented the first results of the seroprevalence test [17], and on May 19^{th}, the Government declared the compulsory use of masks [18], which reinforced the control measures.

In this study, we will do a reconstruction from February 17^{th}, ten days before the first reported case, until May 31^{st}, one day before the AMB entered the Phase One of the de-escalation. We will use **days** as the time units, i.e., we will study the evolution day by day. If at the end of some day t_f all the population is susceptible, recovered or dead, we stop our simulation, since there is no more possibility of infection. In this model, we will consider a homogeneous distribution of the population (without considerations of age or gender) and will not take into account natality and mortality rates, since we will work in a period of few months, so births and natural deaths may be negligible in the model. We will neither consider migratory fluxes for simplicity and due to the control measures imposed to stop the pandemic. Given these assumptions, we present the following model, summarized in a diagram in Figure 6:

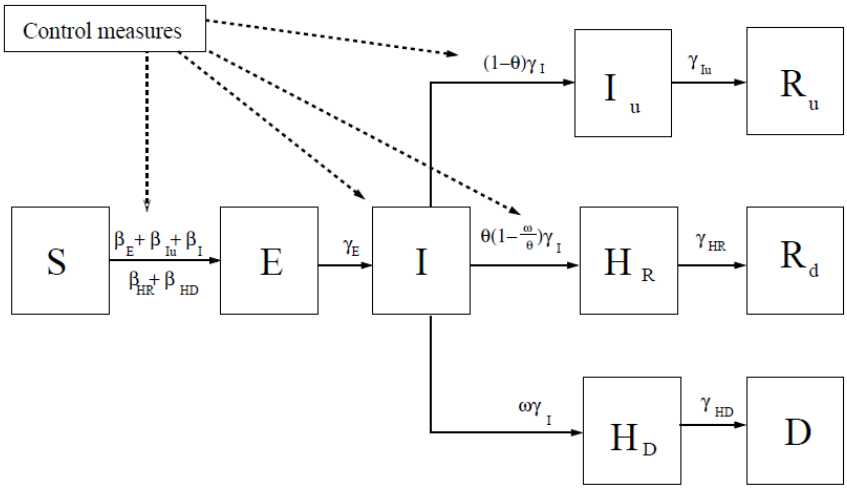


Figure 6: Diagram summarizing the epidemiological model θ-SEIHRD for COVID-19 given by (9).

$$
\begin{cases}
\dfrac{\mathrm{d}S}{\mathrm{d}t}(t) &= -\dfrac{S(t)}{N}\left(m_E(t)\beta_E E(t) + m_I(t)\beta_I I(t) + m_{I_u}(t)\beta_{I_u}(\theta(t))I_u(t)\right) - \\
&\quad -\dfrac{S(t)}{N}\left(m_{H_R}(t)\beta_{H_R}H_R(t) + m_{H_D}(t)\beta_{H_D}H_D(t)\right), \\
\dfrac{\mathrm{d}E}{\mathrm{d}t}(t) &= \dfrac{S(t)}{N}\left(m_E(t)\beta_E E(t) + m_I(t)\beta_I I(t) + m_{I_u}(t)\beta_{I_u}(\theta(t))I_u(t)\right) + \\
&\quad +\dfrac{S(t)}{N}\left(m_{H_R}(t)\beta_{H_R}H_R(t) + m_{H_D}(t)\beta_{H_D}H_D(t)\right) - \gamma_E E(t), \\
\dfrac{\mathrm{d}I}{\mathrm{d}t}(t) &= \gamma_E E(t) - \gamma_I(t)I(t), \\
\dfrac{\mathrm{d}I_u}{\mathrm{d}t}(t) &= (1-\theta(t))\gamma_I I(t) - \gamma_{I_u}I_u(t), \\
\dfrac{\mathrm{d}H_R}{\mathrm{d}t}(t) &= \theta(t)(1-\tfrac{\omega(t)}{\theta(t)})\gamma_I(t)I(t) - \gamma_{H_R}(t)H_R(t), \\
\dfrac{\mathrm{d}H_D}{\mathrm{d}t}(t) &= \omega(t)\gamma_I(t)I(t) - \gamma_{H_D}(t)H_D(t), \\
\dfrac{\mathrm{d}R_d}{\mathrm{d}t}(t) &= \gamma_{H_R}(t)H_R(t), \\
\dfrac{\mathrm{d}R_u}{\mathrm{d}t}(t) &= \gamma_{I_u}(t)I_u(t), \\
\dfrac{\mathrm{d}D}{\mathrm{d}t}(t) &= \gamma_{H_D}(t)H_D(t),
\end{cases}
\tag{9}
$$

Let us notice that the equations for the recovered and dead individuals may be treated separately from the rest, and thus can be computed later using their *explicit* form

$$
R_d(t) = R_d(t_0) + \int_{t_0}^{t} \gamma_{H_R}(s)H_R(s)\ \mathrm{d}s, \quad R_u(t) = R_u(t_0) + \int_{t_0}^{t} \gamma_{I_u}(s)I_u(s)\ \mathrm{d}s, \quad D(t) = D(t_0) + \int_{t_0}^{t} \gamma_{H_d}(s)H_D(s)\ \mathrm{d}s.
$$

Now, we define all the parameters used in system (9):

- $N \in \mathbb{N}$ (number of individuals): population of the AMB before the pandemic.
- $\omega(t) \in [\underline{\omega}, \overline{\omega}]$ (adimensional): infection fatality rate (IFR), proportion of infected people who do not survive the disease. The quantities $\underline{\omega}$ and $\overline{\omega}$ are the minimum and the maximum IFRs, respectively.
- $\theta(t) \in [\overline{\omega}, 1]$ (adimensional): proportion of infected people detected and documented by the authorities. We suppose for simplicity that all deaths due to COVID-19 are detected and

reported, and thus $\theta \geq \overline{\omega}$. The quantity $\omega(t)/\theta(t) \leq 1$ corresponds to the case fatality rate (CFR), the proportion of deaths compared to the detected infections.

- $\beta_E, \beta_I, \beta_{H_R}, \beta_{H_D} \in \mathbb{R}^+$ (day^{-1}): disease effective contact rates of a person in the corresponding state, i.e., the effective contact rate for an infected individual to infect a susceptible person per day.
- $\beta_{I_u}(\theta) \in \mathbb{R}^+$ (day^{-1}): disease effective contact rate of an undetected infectious person in terms of the portion of detected infected individuals.
- $\gamma_E \in (0, +\infty)$ (day^{-1}): transition rate from E to infection I.
- $\gamma_{I_u}(t)$, $\gamma_{H_R}(t)$, $\gamma_{H_D}(t) \in (0, +\infty)$ (day^{-1}): transition rates from I_u, H_R and H_D to R_u, R_d and D, respectively.
- $\gamma_I(t) \in (0, +\infty)$ (day^{-1}): transition rate from infected I to undetected (I_u) or hospitalized (H_R, H_D).
- $m_E(t)$, $m_{I_u}(t)$, $m_I(t)$, $m_{H_R}(t)$, $m_{H_D}(t) \in [0, 1]$ (%): functions representing the efficiency of the control measures applied to each corresponding state: $m_X = 1$ means there are no control measures for state X, and any value $m_X < 1$, since it multiplies β_X, will directly diminish the transmission of the disease from compartment X.

Such as observed with the SIRS model, our system of equations is also biologically consistent and deals with a constant population, which intuitively makes sense since we are also considering the recoveries and deaths.

Theorem 4. *The set*

$$\Omega = \{(S, E, I, I_u, H_R, H_D, R_d, R_u, D) \in [0, +\infty)^9 : S + E + I + I_u + H_R + H_D + R_d + R_u + D = N\}$$

*is **positively invariant** for system (9), given that* $S(0) + E(0) + I(0) + I_u(0) + H_R(0) + H_D(0) + R_d(0) + R_u(0) + D(0) = N$.

Proof. Let Φ_t be the continuous flux of system (9). Then, given an initial condition $\boldsymbol{X} \in [0, +\infty)$, corresponding $\boldsymbol{X}$ to $(S, E, I, I_u, H_R, H_D, R_d, R_u, D)$, let us see that

$$\Phi_t(\boldsymbol{X}) \in [0, +\infty)^9, \, \forall t \geq 0.$$

To do that, let $\boldsymbol{X}(\bar{t}) \in \Omega$, for some $\bar{t} \geq 0$. Then, it is straightforward to deduce that

$$X_i(\bar{t}) = 0 \implies \frac{\mathrm{d}X_i}{\mathrm{d}t}(\bar{t}) \geq 0, \, \forall i = 1, ..., 9.$$

Hence, for any initial condition $\boldsymbol{X}(0) \in \Omega$, given that our flow Φ_t is continuous, if any of the derivatives is negative, there may exist times t_i, $i = 1, ..., 9$ such that $X_i(t_i) = 0$. Without loss of generality, we suppose $t_1 \leq ... \leq t_9$. Therefore,

$$\Phi_{t_1}(\boldsymbol{X}_0) = (0, X_2(t_1), ..., X_9(t_1)) \in \{0\} \times [0, \infty)^8.$$

Now, since $\frac{\mathrm{d}X_1}{\mathrm{d}t}(t_1) \geq 0$ as stated before, X_1 will either increase or remain being 0, and hence will remain positive or null.

This is valid for the 9 compartments. Thus, we deduce that Ω is a positive invariant set for our continuous flow $\Phi_{t\geq 0}$.

On the other hand, for any initial solution $(S_0, ..., D_0) \in \Omega$, $S_0 + ... + D_0 = N$. Adding up the 9 equations of our system, we obtain

$$\frac{\mathrm{d}S}{\mathrm{d}t}(t) + ... + \frac{\mathrm{d}D}{\mathrm{d}t}(t) = 0, \, \forall t \geq 0.$$

This is a first integral that implies the sum of all the compartments remains constant, i.e., $S(t) + ... + D(t) = S_0 + ... + D_0 = N$, which leads us to conclude that Ω is positively invariant with respect to the flow of system (9). □

This result will be relevant in order to study the stability of the system. Next, we compute its basic and effective reproduction numbers R_0 and R_t, respectively.

3.2 Computation of the reproductive numbers R_0 and R_t

For this task, we start computing R_0 by using the *Next Generation Matrix* as done previously.

In this case, we have the following infected states: E, I_d, I_u, H_R, H_D. Thus, $m = 5$. We sort our compartmental vector as $\boldsymbol{C} = (E, I_d, I_u, H_R, H_D, S, R_d, R_u, D)$ and define our functions $F(\boldsymbol{C})$ and $V(\boldsymbol{C})$ as follows:

$$F(\boldsymbol{C}) = \begin{pmatrix} \frac{S(t)}{N}\mathrm{INF}(t) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, V(\boldsymbol{C}) = \begin{pmatrix} \gamma_E E(t) \\ -\gamma_E E(t) + \gamma_I I(t) \\ -(1-\theta)\gamma_I I(t) + \gamma_{I_u} I_u(t) \\ -\theta(1-\frac{\omega}{\theta})\gamma_I I(t) + \gamma_{H_R} H_R(t) \\ -\omega\gamma_I I(t) + \gamma_{H_D} H_D(t) \\ \frac{S(t)}{N}\mathrm{INF}(t) \\ -\gamma_{H_R}(t) H_R(t) \\ -\gamma_{I_u}(t) I_u(t) \\ -\gamma_{H_D}(t) H_D(t) \end{pmatrix}, \tag{10}$$

where

$$\mathrm{INF}(t) = \beta_E E(t) + \beta_{I_u} I_u(t) + \beta_I I(t) + \beta_{H_R} H_R(t) + \beta_{H_D} H_D(t).$$

Notice that, to simplify notations, each parameter function corresponds to its value at t_0.

The next step is to compute the Jacobian matrices of these functions:

$$DF(\boldsymbol{C}) = \begin{pmatrix} \beta_E \frac{S(t)}{N} & \beta_I \frac{S(t)}{N} & \beta_{I_u} \frac{S(t)}{N} & \beta_{H_R} \frac{S(t)}{N} & \beta_{H_D} \frac{S(t)}{N} & \frac{\text{INF}(t)}{N} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$DV(\boldsymbol{C}) = \begin{pmatrix} \gamma_E & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\gamma_E & \gamma_I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -(1-\theta)\gamma_I & \gamma_{I_u} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -(\theta-\omega)\gamma_I & 0 & \gamma_{H_R} & 0 & 0 & 0 & 0 & 0 \\ 0 & -\omega\gamma_I & 0 & 0 & \gamma_{H_D} & 0 & 0 & 0 & 0 \\ \beta_E \frac{S(t)}{N} & \beta_I \frac{S(t)}{N} & \beta_{I_u} \frac{S(t)}{N} & \beta_{H_R} \frac{S(t)}{N} & \beta_{H_D} \frac{S(t)}{N} & \frac{\text{INF}(t)}{N} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\gamma_{H_R} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\gamma_{I_u} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\gamma_{H_D} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

To continue on the computations of R_0, we need to determine the disease-free equilibrium **where all the individuals are susceptible**. It is clear that, since we have a constant population of N people, our desired point will be

$$\boldsymbol{C}_f = (0, 0, 0, 0, 0, N, 0, 0, 0),$$

which satisfies the condition for being an equilibrium.

Therefore, we can compute our new Jacobian matrices associated to $\boldsymbol{C}_f$, obtaining

$$DF(\boldsymbol{C}) = \begin{pmatrix} \beta_E & \beta_I & \beta_{I_u} & \beta_{H_R} & \beta_{H_D} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$DV(\boldsymbol{C}) = \begin{pmatrix} \gamma_E & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\gamma_E & \gamma_I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -(1-\theta)\gamma_I & \gamma_{I_u} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\theta(1-\frac{\omega}{\theta})\gamma_I & 0 & \gamma_{H_R} & 0 & 0 & 0 & 0 & 0 \\ 0 & -\omega\gamma_I & 0 & 0 & \gamma_{H_D} & 0 & 0 & 0 & 0 \\ \beta_E & \beta_I & \beta_{I_u} & \beta_{H_R} & \beta_{H_D} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\gamma_{H_R} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\gamma_{I_u} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\gamma_{H_D} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since $m = 5$, we need to extract the first 5×5 blocks, and, therefore,

$$\mathcal{F}\mathcal{V}^{-1} = \begin{pmatrix} \beta_E & \beta_I & \beta_{I_u} & \beta_{H_R} & \beta_{H_D} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\gamma_E & 0 & 0 & 0 & 0 \\ 1/\gamma_I & 1/\gamma_I & 0 & 0 & 0 \\ (1-\theta)/\gamma_{I_u} & (1-\theta)/\gamma_{I_u} & 1/\gamma_{I_u} & 0 & 0 \\ (\theta-\omega)/\gamma_{H_R} & (\theta-\omega)/\gamma_{H_R} & 0 & 1/\gamma_{H_R} & 0 \\ \omega/\gamma_{H_D} & \omega/\gamma_{H_D} & 0 & 0 & 1/\gamma_{H_D} \end{pmatrix} =$$

$$= \begin{pmatrix} \mu_E + \mu_I + \mu_{H_R}(\theta-\omega) + \mu_{I_u}(1-\theta) + \mu_{H_D}\omega & \mu_I + \mu_{H_R}(\theta-\omega) + \mu_{I_u}(1-\theta) + \mu_{H_D}\omega & \mu_{I_u} & \mu_{H_R} & \mu_{H_D} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence,

$$R_0 = \mu_E + \mu_I + \mu_{H_R}(\theta-\omega) + \mu_{I_u}(1-\theta) + \mu_{H_D}\omega,$$

where we denote $\mu_X = \beta_X/\gamma_X$. Notice that $R_0 \geq 0$, since our parameters are non-negative and $\overline{\omega} \leq \theta \leq 1$.

Followingly, given that the system is non-autonomous, depending on the evolution of our parameters (specially our control measures $m_X(t)$), this R_0 may not be valid after some time, since it is defined on a whole susceptible population; for this reason, it is useful computing in this case also the effective reproduction number R_t, which should be similar to R_0 (remember we already mentioned that, for an autonomous system, they are related as $R_t = (S(t)/N)R_0$).

If we repeat the same process, in this case the point on which we evaluate our matrices will not be $\boldsymbol{C}_f$ but

$$\boldsymbol{C}_t = (E_t, I_t, I_{u,t}, H_{R,t}, H_{D,t}, R_{d,t}, R_{u,t}, D_t),$$

i.e., the individuals in each compartment at the instant t. Focusing on (10), we will have the same functions but with the difference that, in this case,

$$\text{INF(t)} = m_E(t)\beta_E E(t) + m_I(t)\beta_I I(t) + m_{I_u}(t)\beta_{I_u}(\theta)I_u(t) + m_{H_R}(t)\beta_{H_R}H_R(t) + m_{H_D}\beta_{H_D}H_D(t).$$

Performing the same computations as before, we get

$$\mathcal{F} = \begin{pmatrix} m_E(t)\beta_E S_t/N & m_I(t)\beta_I S_t/N & m_{I_u}\beta_{I_u}(\theta)S_t/N & m_{H_R}(t)\beta_{H_R}S_t/N & m_{H_D}(t)\beta_{H_D}S_t/N \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathcal{V} = \begin{pmatrix} \gamma_E & 0 & 0 & 0 & 0 \\ -\gamma_E & \gamma_I & 0 & 0 & 0 \\ 0 & -(1-\theta(t))\gamma_I & \gamma_{I_u} & 0 & 0 \\ 0 & -(\theta(t)-\omega(t))\gamma_I & 0 & \gamma_{H_R} & 0 \\ 0 & -\omega(t)\gamma_I & 0 & 0 & \gamma_{H_D} \end{pmatrix},$$

from which we obtain

$$\begin{aligned} R_t = \frac{S_t}{N} \left(m_E(t)\mu_E + m_I(t)\mu_I + m_{I_u}(t)\mu_{I_u}(\theta(t))(1-\theta(t)) + \right. \\ \left. + m_{H_R}(t)\mu_{H_R}(\theta(t)-\omega(t)) + m_{H_D}(t)\mu_{H_D}\omega(t)\right). \end{aligned} \tag{11}$$

We note that it is very similar to our R_0, but depends on the temporal evolution of our parameters. However, R_t is also non-negative, for every $t \geq 0$.

Let us now compute the rest of the equilibria and see their stability with respect to our reproduction numbers.

3.3 Equilibria and stability

To start with, notice that there is **no endemic equilibrium**. In fact, any equilibrium needs to fulfill $\mathrm{d}S/\,\mathrm{d}t = 0$, i.e.,

$$\frac{S}{N}(\beta_E E + \beta_I I + \beta_{I_u} I_u + \beta_{H_R} H_R + \beta_{H_D} H_D) = 0.$$

Here, we find two options: $S = 0$ or $\mathrm{INF} = 0$. Since $\beta_X > 0$ and each compartment is also non-negative, $\mathrm{INF} = 0$ would mean that each infected state is equal to zero. Therefore, if there was an endemic equilibrium, we would need $S = 0$, this is, the whole population has been infected. Consequently, it would be obtained

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{S}{N}\mathrm{INF} - \gamma_E E = -\gamma_E E = 0,$$

which implies $E = 0$. It is straightforward seeing that this also implies all the infected compartments are empty, and, hence, **no endemic equilibrium is possible**. In fact, this makes sense, given that our very last states are recovery (R_u and R_d) and death (D), and it is naturally clear that every infected individual will eventually end up in some of these states. Moreover, it is worth to remark they cannot be reinfected since our model is not circular and there is no incoming flux for the S compartment.

Thus, our equilibria will be of the form

$$\boldsymbol{C} = (0,0,0,0,0,S^*,R_d^*,R_u^*,D^*), \quad S^* + R_d^* + R_u^* + D^* = N. \tag{12}$$

This would mean that we have, *a priori*, a 3-dimensional space of equilibria. Nevertheless, it is not true since we have some *hidden* constraints due to ω and θ. Indeed,

$$\mathrm{IFR}(t) = \frac{H_D + D}{I + I_u + H_R + H_D + R_u + R_d + D}, \quad \mathrm{CFR}(t) = \frac{I + H_R + R_d + H_D + D}{I + I_u + H_R + H_D + R_u + R_d + D}, \tag{13}$$

where CFR(t) and IFR(t) are the **instantaneous** case and infection fatality rates, respectively. Nevertheless, notice that these rates are *a priori* **unknowns**.

Considering now that $I^* = I_u^* = H_R^* = H_D^* = 0$, we would obtain two constraining relations between R_u^*, R_d^* and D^*. Hence, it seems that we have a 4-dimensional hyperspace with three constraints ((12) and both from (13)), which would lead us to a 1-dimensional manifold of equilibria. However, notice that two of our constraints depend on the time instant we are at, since the instantaneous IFR and CFR may change with the improvement or relaxation of the control measures, for instance. This implies that we have one more dimension (time), and thus in reality this would be the **equilibrium 2-dimensional manifold**

$$C = \{\boldsymbol{C} = (0,0,0,0,0,S^*,R_d^*,R_u^*,D^*) \in \{0\}^5 \times [0,+\infty)^4 : \boldsymbol{C} \text{ satisfies (12) and (13)}, \forall t \geq 0\} \subset \mathbb{R}^5.$$

Theorem 5. *The equilibrium* $2D$*-manifold of system (9)*

$$C = \{\boldsymbol{C} = (0,0,0,0,0,S^*,R_d^*,R_u^*,D^*) \in \{0\}^5 \times [0,+\infty)^4 : \boldsymbol{C} \textit{ satisfies (12) and (13)}, \forall t \geq 0\} \subset \mathbb{R}^5$$

is a GAS manifold for any $R_0 > 0$.

Proof. Following the procedure used in the proof of Theorem 2, in this case we also have $\mathcal{F}, \mathcal{V}^{-1} > 0$. Restricting our system as previously justified, we have $\boldsymbol{x} = (E, I, I_u, H_R, H_D)$ and $y = S$, and our function $f(\boldsymbol{x}, y)$ is

$$f(\boldsymbol{x}, y) = (\mathcal{F} - \mathcal{V})\boldsymbol{x} - F + V = \left(\beta_E E + \left(1 + \frac{S}{N}\right)\text{INF}, 0, 0, 0, 0\right)^T,$$

which is clearly non-negative for every $(\boldsymbol{x}, y) \in \Omega$ (here we make an abuse of notation naming Ω the previous Ω set restricted to $(\boldsymbol{x}, y)$). Hence, we rewrite our system for the infected compartments as

$$\dot{\boldsymbol{x}} = (\mathcal{F} - \mathcal{V})\boldsymbol{x} - f(\boldsymbol{x}, y).$$

Now, notice that, if R_0 is an eigenvalue for $\mathcal{F}\mathcal{V}^{-1}$, so it is for $\mathcal{V}^{-1}\mathcal{F}$, since

$$\det(\mathcal{F}\mathcal{V}^{-1} - R_0 I_5) = 0 \implies \det(\mathcal{F} - R_0\mathcal{V}) = 0 \implies \det(\mathcal{V}^{-1}\mathcal{F} - R_0 I_5) = 0.$$

Thus, we can find $\nu^T \geq 0$ the left eigenvector of $\mathcal{V}^{-1}\mathcal{F}$ for the eigenvalue R_0:

$$\nu^T\mathcal{V}^{-1}\mathcal{F} = \nu^T \begin{pmatrix} \mu_E & \beta_I/\gamma_E & \beta_{I_u}/\gamma_E & \beta_{H_R}/\gamma_E & \beta_{H_D}/\gamma_E \\ \beta_E/\gamma_I & \mu_I & \beta_{I_u}/\gamma_I & \beta_{H_R}/\gamma_I & \beta_{H_D}/\gamma_I \\ (1-\theta)\beta_E/\gamma_{I_u} & (1-\theta)\beta_I/\gamma_{I_u} & (1-\theta)\mu_{I_u} & (1-\theta)\beta_{H_R}/\gamma_{I_u} & (1-\theta)\beta_{H_D}/\gamma_{I_u} \\ (\theta-\omega)\beta_E/\gamma_{H_R} & (\theta-\omega)\beta_I/\gamma_{H_R} & (\theta-\omega)\beta_{I_u}/\gamma_{H_R} & (\theta-\omega)\mu_{H_R} & (\theta-\omega)\beta_{H_D}/\gamma_{H_R} \\ \omega\beta_E/\gamma_{H_D} & \omega\beta_I/\gamma_{H_D} & \omega\beta_{I_u}/\gamma_{H_D} & \omega\beta_{H_R}/\gamma_{H_D} & \omega\mu_{H_D} \end{pmatrix} = R_0\nu^T.$$

If we look only at the first component and recall R_0 was

$$R_0 = \mu_E + \theta\mu_I + (1-\theta)\mu_{I_u} + (\theta-\omega)\mu_{H_R} + \omega\mu_{H_D},$$

we can easily deduce

$$\nu^T = (1, \beta_I/\beta_E, \beta_{I_u}/\beta_E, \beta_{H_R}/\beta_E, \beta_{H_D}/\beta_E).$$

These are a little bit cumbersome computations, but it can be proved with some symbolic program that this equality with this ν is also fulfilled for the other four components (which makes sense, since R_0 is an eigenvalue and we have already found a possible eigenvector associated to it).

Then, attending to Theorem 3,

$$Q(\boldsymbol{x}) = \nu^T \mathcal{V}^{-1} \boldsymbol{x}$$

is a Lyapunov-type function for our system. But, do we obtain from here global asymptotic stability? We can see explicitly who this Q is:

$$Q(\boldsymbol{x}) = \nu^T \mathcal{V}^{-1} \begin{pmatrix} E \\ I \\ I_u \\ H_R \\ H_D \end{pmatrix} = \frac{E}{\gamma_E} + \frac{\mu_I}{\beta_E}(E+I) + \frac{\mu_{I_u}}{\beta_E}((1-\theta)(E+I)+I_u) + \frac{\mu_{H_R}}{\beta_E}((\theta-\omega)(E+I)+H_R) + \frac{\mu_{H_D}}{\beta_E}(\omega(E+I)+H_D).$$

Such as we saw in the proof of Theorem 3,

$$\dot{Q}(\boldsymbol{x}) = \nu^T (R_0 - 1)\boldsymbol{x} - \nu^T \mathcal{V}^{-1} f(\boldsymbol{x}, y).$$

Hence, we may proceed to prove the global stability by using again the LaSalle's principle again:

Since we are computing at $t = 0$, all our parameters are constant and this, with the fact that (E, I, I_u, H_R, H_D) are bounded means that there exists some constant $c \in \mathbb{R}^+$ such that $Q(\boldsymbol{x}) < c$ in Ω. Now it only remains to see who Γ is. Let $R_0 \leq 1$, then we need

$$f(\boldsymbol{x}, y) = 0 \implies \beta_E E + \left(1 + \frac{S}{N}\right) \text{INF} = 0.$$

Given that $\boldsymbol{x}, y \geq 0$ and our parameters are also positive, we need

$$\text{INF} = 0 \implies E = I = I_u = H_R = H_D = 0,$$

and therefore our maximal invariant subset is

$$\Gamma_0 = \{(0, 0, 0, 0, 0, S, R_d, R_u, D) \in \Omega\},$$

from which we obtain, as expected, that C is globally asymptotically stable for $R_0 \leq 1$.

What happens for $R_0 > 1$? Well... It is also GAS! Let us have a closer look at $\nu^T \mathcal{V}^{-1} f(\boldsymbol{x}, y)$:

$$\begin{aligned} \nu^T \mathcal{V}^{-1} f(\boldsymbol{x}, y) &= \nu^T \mathcal{V}^{-1} \boldsymbol{e}_1 \left(\beta_E E + \left(1 + \frac{S}{N}\right) \text{INF}\right) = \\ &= \frac{1}{\beta_E} (\beta_E, \beta_I, \beta_{I_u}, \beta_{H_R}, \beta_{H_D}) \begin{pmatrix} 1/\gamma_E \\ 1/\gamma_I \\ (1-\theta)/\gamma_{I_u} \\ (\theta-\omega)/\gamma_{H_R} \\ \omega/\gamma_{H_D} \end{pmatrix} \left(\beta_E E + \left(1 + \frac{S}{N}\right) \text{INF}\right) = \\ &= \frac{1}{\beta_E} R_0 \left(\beta_E E + \left(1 + \frac{S}{N}\right) \text{INF}\right), \end{aligned}$$

where $\boldsymbol{e}_1$ is the first vector of the canonical basis of $\mathbb{R}^5$. Besides, recall that

$$\nu^T(R_0-1)\boldsymbol{x} = (R_0-1)\nu^T\boldsymbol{x} = \frac{1}{\beta_E}(R_0-1)\text{INF}.$$

Therefore,

$$\dot{Q}(\boldsymbol{x}) = \nu^T(R_0-1)\boldsymbol{x} - \nu^T\mathcal{V}^{-1}f(\boldsymbol{x},y) = \frac{1}{\beta_E}\left(-\text{INF} - R_0\beta_E E - R_0\frac{S}{N}\text{INF}\right) < 0,$$

for all $\boldsymbol{x} \neq \boldsymbol{0}$.

Consequently, LaSalle's principle is also applicable in this case and C results to be a **globally asymptotically 2D-manifold** for any value of $R_0 > 0$. □

Given this result, one may think this is not contributing with new information, since it is intuitively clear, because of the structure of our model, that every individual will eventually be either recovered, dead or will keep on being susceptible. However, it is important to check that the mathematical results are consistent with the biological behaviour, and obtaining a different result would have meant we are working with an invalid model.

Finally, our following task would have been observing if we can tend to a point with susceptible population or if $S = 0$ is the real GAS equilibrium. In first place, it is arguably clear that $S = N$ is **unstable** due to the non-circularity of our model. Nevertheless, after some failed results, we have not found a theoretical way to determine the precise fate of any concrete point starting with both susceptible and infected population.

3.4 Numerical study

3.4.1 Context and considerations

Now, we present some numerical results aiming to reproduce the observed behaviour of the pandemic in the Metropolitan Area of Barcelona (say, AMB due to its initials in Catalan). The considered data have been recompiled from the retrospection dispensed by the *Departament de Salut, Generalitat de Catalunya* on June 21^{st} [19], which can be found at `https://govern.cat/salapremsa/notes-premsa/386043/comunicat-del-departament-salut`, and the simulations have been run from February 17^{th}, ten days before the report of the first case. Nevertheless, there have been recently discovered rests of SARS-CoV-2 in Catalan waste waters from January 15^{th} [20], i.e., it will be necessary to suppose some amount of initial cases on February 17^{th}, even though they were not reported; this lack of reporting could be due to the fact that it is estimated there are around 80% of asymptomatic individuals or with low symptoms[1], added to the fact that, given the lack of knowledge about the disease at that period, hospitals may have done wrong diagnostics cataloguing it as, for example, influenza. Therefore, it will be very important to estimate properly the initial conditions in order to fit duly the reported evolution.

[1]This can be estimated by using the results of the seroprevalence test performed on May 13^{th}: it showed that 7.1% of the population in Barcelona was infected, which compared to the reported cases of that day leads to a detection of $\sim 20\%$.

Before discussing the estimation of the parameters involved in our model, we should have some considerations first with respect to this study:

- First of all, one must take into account that this is a new disease and many studies are currently being performed to obtain more information about it. It has been observed, for example, that the lethality may substantially vary from one region to another, depending on the sanitary services, mean age of the population, etc. Hence, in this work, we pretend to estimate the parameters from the recompiled data, by considering existing studies (e.g., the study by B. Ivorra, A. M. Ramos et al. [21]) so as to have a first intuition.
- Thus, this study only pretends to be a first approach to know better the evolution of the pandemic in the AMB. Although some types of *predictions* will be shown, they just aim to expose an intuition of what could happen with the relaxation of the control measures, but do not pretend to fit with the actual evolution during the de-escalation.
- Besides, it is well-known the lack of protocols for recompiling the data and the continuous corrections that are being made [22]. This is, although we use official data to obtain the parameters, these parameters do not necessarily correspond to the real situation; therefore, the results for our uncertainty compartments (i.e., I_u and R_u) have to be analyzed carefully and under some doubt. As aforementioned, this study is just a first approach whose main objective is reproducing the data to illustrate how a mathematical epidemiological model works for a real and complex case.
- Finally, this study leaves an open way for future improvements, such as taking into account more compartments (improvements that are already being made by the team B. Ivorra, M. R. Ferrández, A. M. Ramos and M. Vela-Pérez), adding a vaccine, automatizing the fitting processes, etc.

3.4.2 Estimating the parameters

In the following lines, we present some brief explanations on the estimations of the parameters in order to obtain the presented numerical results. We have made the calibration by hand due to the high computational cost supposed by an automatization, and using trial-and-error analyzing the obtained figures.

- **Proportion of detected infections $\theta(t)$:** This is a very important parameter in order to know the real magnitude of the pandemic, but is also very difficult to estimate since it is based on unknown data. In our case, the results of the seroprevalence test exposed on May 13^{th} have been specially helpful [17]. Using this, we have estimated $\theta(t)$ in the following way:

  In Figure 7, we present an inference of the evolution of the prevalence in the AMB. Since the increasing of these diseases is exponential, we have supposed an exponential growth of the prevalence before the establishment of the first control measures on March 15^{th}; in fact, this exponential growth is kept some more days due to the initial disobedience to the alarm state, until March 28^{th}, when the workers of no-essential services were also obligated to stay

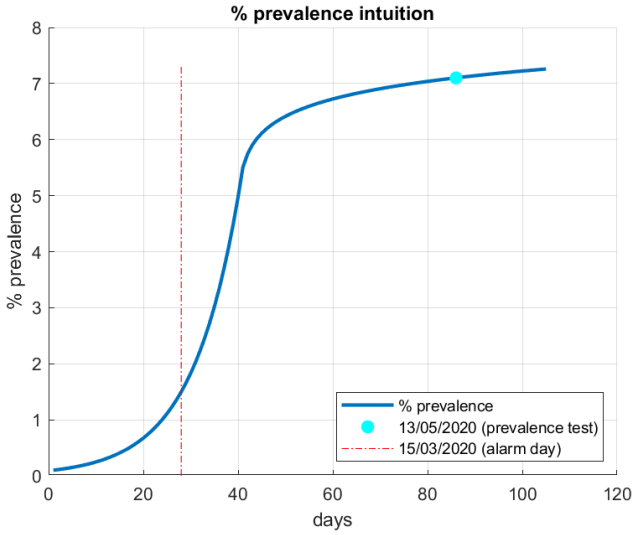


Figure 7: Intuition of the evolution of the prevalence among the AMB population based on the seroprevalence test carried out on May 13^{th}.

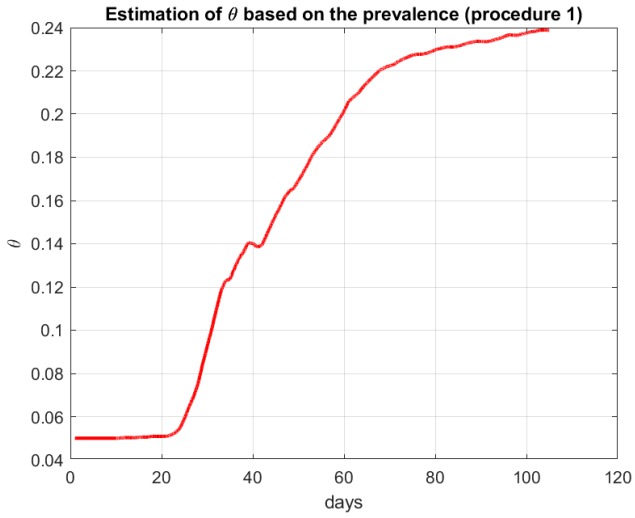


Figure 8: Estimation of θ based on the prevalence and the recompiled data.

home. Then, it is supposed a huge decreasing on the growth due to this state, inferring the prevalence will be maintained at $\sim 7.1\%$.

Given this, in Figure 8 we present the estimate for θ: since $\theta(t) =$ (cumulative detected infected)(t)/(cumulative total infected)(t), we have estimated it as

$$\theta(t) = \frac{\text{detected}(t)}{N(\text{prevalence}(t))},$$

where "detected" refers to the daily accumulated values of the known detections communicated by the *Departament de Salut*, t is counted in days, prevalence$(t) \in [0, 1]$ and N is the total population of the AMB - we have considered $N = 3,239,337$. Besides, we have added 0.05 to our result, basing ourselves on the study for Madrid [21].

Notice that, with this inference, no predictions can be made, given that we are using the communicated data, unless we analyze the possible future behaviour of the pandemic and the detections. For that, in Figure 9 we have supposed $\theta(t)$ as a linear function basing ourselves on the previous data and taking into account the linear approximation made for China in [13].

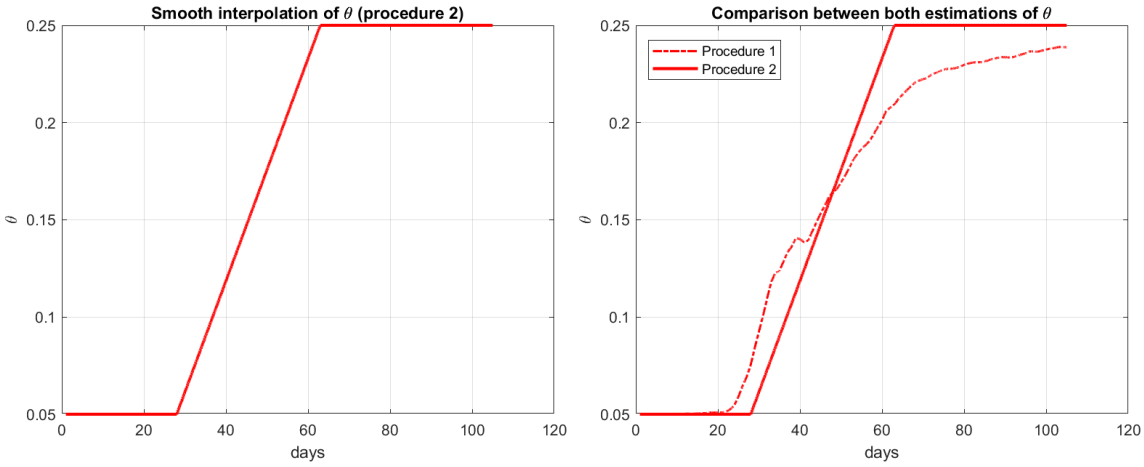


Figure 9: On the left, a linear estimation for θ derived from the first procedure in Figure 8 and based on the study for China in [13]. On the right, a comparison between both estimations.

This way,

$$\theta(t) = \begin{cases} 0.05,\ t \in [\text{February } 17^{th},\ \text{March } 15^{th}), \\ \text{linear continuous},\ t \in [\text{March } 15^{th},\ \text{April } 19^{th}), \\ 0.25,\ t \geq \text{April } 19^{th}. \end{cases}$$

It is also shown a comparison between the resulting estimations using both methods.

- **Infection fatality rate $\omega(t)$:** To obtain this value, we are going to use again our data and the results for $\theta(t)$. In fact, only using the reported data for positive cases and deaths, we can compute the CFR by dividing the daily cumulative positive cases times the daily cumulative deaths, which would correspond to $\omega(t)/\theta(t)$. However, this rate should consider a **delay** on the deceases, since a positive case may not die until 2 or 3 weeks after the detection. Consequently, we have supposed a delay of **15 days**, which can be considered with our data, since our computations are until May 31^{st} and we have further reports. In Figure 10, we present an intuition for ω, estimated as

$$\omega(t) = \frac{\text{deaths}(t+15)}{N(\text{prevalence}(t))}.$$

Such as previously, we have added 0.001 to consider some initial mortality on the initial unreported cases.

Observe that we are dealing with high values of ω, of the order of a 4%. Besides, we appreciate an abrupt growth generating a first peak - this peak is close to March 15^{th} and possibly corresponds to a saturation of the sanitary system. In Figure 11, we suppose a smoother inference correcting those two peaks and present a comparison between both estimations. In Section 3.4.3, we will check the accuracy of these second interpolations for θ and ω.

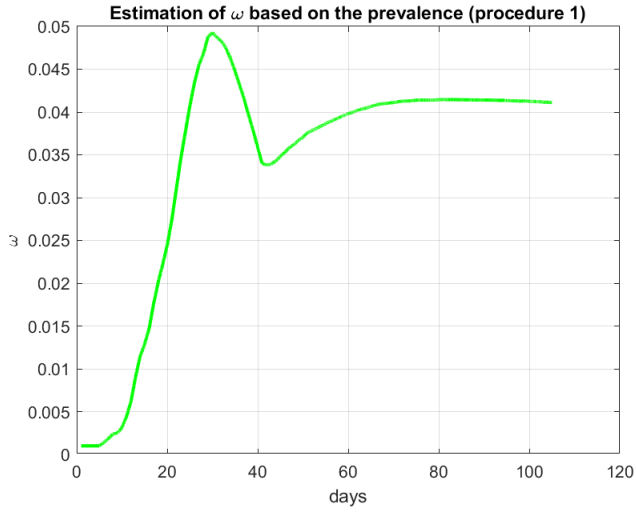


Figure 10: Estimation of ω based on the prevalence and the recompiled data, supposing a delay of 15 days before the decease.

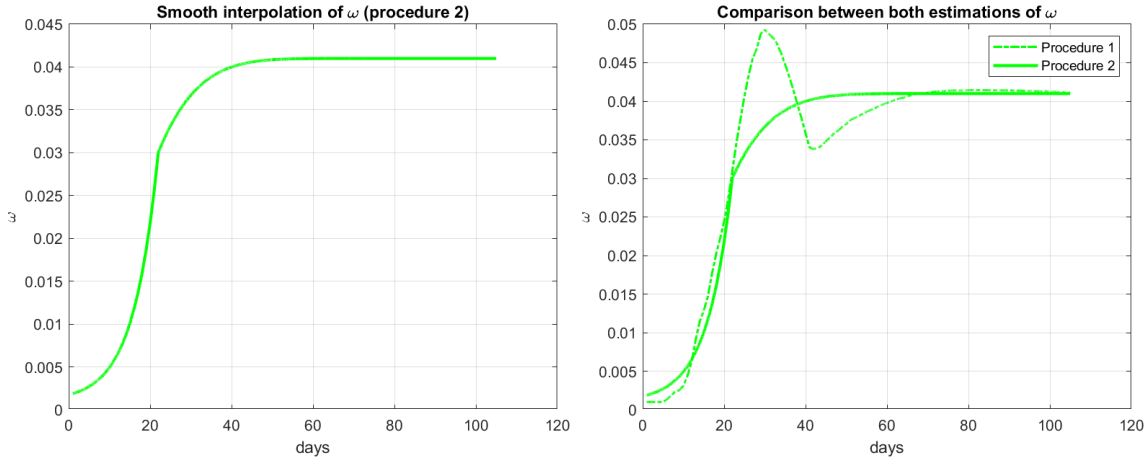


Figure 11: On the left, a smooth estimation of ω based on the one in Figure 10. On the right, a comparison between the estimations of both procedures.

Observe that, such as we asked in Section 3.1, θ fulfills the condition

$$\theta(t) \geq \overline{\omega}$$

in both estimations.

- **Control measures** m_X**:** To estimate the functions of these measures, we consider [13]. To do that, we revise again some important dates:
  - **March** 15^{th}**:** First day of the alarm state in Spain, 28 days after February 17^{th}.
  - **March** 28^{th}**:** No-essential services are stopped for their workers to stay home, 41 days after February 17^{th}.
  - **April** 12^{th}**:** Workers of no-essential services are allowed to go back to work, 56 days after February 17^{th}.

- **April 26^{th}:** Kids are allowed to go out with their parents, respecting the social distancing, 70 days after February 17^{th}. Besides, on May 2^{nd} people are allowed to exercise on their own and during some determined hours.
- **May 19^{th}:** It becomes compulsory wearing a mask outside, 93 days after February 17^{th}.

From here, we define our control measure functions as follows:

$$m_X(t) = \begin{cases} 1, & t \in [\text{February } 17^{th}, \text{March } 15^{th}), \\ (1 - m_{X,1})\ \exp(-\kappa_1(t-28)) + m_{X,1}, & t \in [\text{March } 15^{th}, \text{March } 28^{th}), \\ (m_{X,1} - m_{X,2})\ \exp(-\kappa_2(t-41)) + m_{X,2}, & t \in [\text{March } 28^{th}, \text{April } 12^{th}), \\ (m_{X,2} - m_{X,3})\ \exp(-\kappa_3(t-56)) + m_{X,3}, & t \in [\text{April } 12^{th}, \text{April } 26^{th}), \\ (m_{X,3} - m_{X,4})\ \exp(-\kappa_4(t-70)) + m_{X,4}, & t \in [\text{April } 26^{th}, \text{May } 19^{th}), \\ (m_{X,4} - m_{X,5})\ \exp(-\kappa_5(t-93)) + m_{X,5}, & t \geq \text{May } 19^{th}, \end{cases}$$

where $m_{X,i} \in [0,1]$ represent the maximum effectiveness of each control measure and $\kappa_i \in [0, 0.2]$ is a fitting parameter. For China, the results were fit with the greatest $\kappa = 0.2$ and the smallest $m = 0$, i.e., the most effective measures for the given interval. However, China is a country that has suffered from many epidemics and both the Government and its citizens were more prepared to confront such a situation; on the contrary, in Spain did not exist such protocols and, therefore, we cannot assume such effective measures. Besides, in [13], the authors consider the same m for each compartment; nevertheless, we have divided it in two groups: same control measures for (I, H_R, H_D) and another different for (E, I_u), due to the false sensation of security these last compartments may perceive when presenting no symptoms.

After studying the implemented measures, their effectiveness (such as the reduction of infection due to masks) and doing some numerical tests, we have supposed the following:

- $m_{I,1} = 0.28$, $m_{E,1} = m_{I,1} + 0.05$, $\kappa_1 = 0.1$,
- $m_{I,2} = 0.11$, $m_{E,2} = m_{I,2} + 0.05$, $\kappa_2 = 0.2$,
- $m_{I,3} = 0.115$, $m_{E,3} = m_{I,3} + 0.05$, $\kappa_3 = 0.2$,
- $m_{I,4} = 0.11$, $m_{E,4} = m_{I,4} + 0.05$, $\kappa_4 = 0.2$,
- $m_{I,5} = 0.08$, $m_{E,5} = m_{I,5} + 0.05$, $\kappa_5 = 0.2$.

Observe that we have considered a little growth from $m_{I,2}$ to $m_{I,3}$, given that the permission for going back to work initially supposed some relaxation of the control measures.

- **Contact rates β_X:** Again, basing ourselves on [13], we define

$$\beta_E = C_E\beta_I, \quad \beta_{H_D}(t) = \beta_{H_R}(t) = C_H(t)\beta_I, \quad \beta_{I_u}(\theta) = C_u\beta_I + \frac{\beta_I(1-C_u)}{1-\omega(t)}(1-\theta),$$

where C_E, C_u are some constants in $[0,1]$ and $C_H(t) \in [0,1]$. for all $t \geq 0$, The expression for $\beta_{I_u}(\theta)$ comes from the following linear relation:

$$\beta_{I_u} = \beta_I \frac{1-\theta}{1-\omega} + \underline{\beta_{I_u}} \frac{\theta - \omega}{1-\omega},$$

where $\underline{\beta_{I_u}} = C_u\beta_I$ is the minimum transmission rate for this compartment. This is, the more cases are detected ($\theta \geq \omega$), the smaller this rate is.

On the other hand, from the aforementioned paper and performing many numerical tests, we have calibrated $\beta_I = 0.1925$, $C_E = 0.25$ and $C_u = 0.4$. Besides, it is estimated that $\sim 10,000$ healthcare workers had been infected in Catalunya until May 21^{st} [23]. On May 20^{th}, there were reported $63,259$ positive cases in Catalunya by the *Departament de Salut*, which means over a 16% of these cases were healthcare workers. Then, from some point advanced in time, we ask the next relation to be fulfilled between the expected hospital cases and the expected total cases:

$$\frac{C_H(t)\beta_I\left((\theta-\omega)m_{H_R}/\gamma_{H_R}+\omega m_{H_D}/\gamma_{H_D}\right)}{C_H(t)\beta_I\left((\theta-\omega)m_{H_R}/\gamma_{H_R}+\omega m_{H_D}/\gamma_{H_D}\right)+m_E\mu_E+m_I\mu_I+(1-\theta)m_{I_u}\mu_{I_u}(\theta)}\approx 0.16,$$

recalling that we defined $\mu_X = \beta_X/\gamma_X$. The expected cases for each compartment may be computed as in our explanation of the estimation of R_0 in Section 2.2. Hence, we estimate

$$C_H(t) = 0.1905\frac{(m_I\mu_I+m_E\mu_E+(1-\theta)m_{I_u}\mu_{I_u}(\theta))}{\beta_I\theta\left((1-\omega/\theta)m_{H_R}/\gamma_{H_R}+(\omega/\theta)m_{H_D}/\gamma_{H_D}\right)}.$$

Nevertheless, if $C_H(t) > 1$, for some t, we naturally set $C_H(t) = 1$.

- **Transition rates γ_X:** The quantities $\gamma_X^{-1} = d_X$ correspond to the (mean) number of days of an individual in state X. It has been analyzed (see [24]) that symptoms after exposure are estimated to appear between the first 5 or 6 days. Besides, an individual with mild or medium symptoms needs in average 14 days to recover, while an infected individual with a severe infection needs from three to six weeks to completely recover. From here, and estimating that an individual will be or not detected in 6.1 days in mean, we set

$$d_E = 5.5, \quad d_I = 6.1, \quad d_{I_u} = d_{H_R} = 14 - d_I, \quad d_{H_D} = d_{H_R} + \delta_D,$$

where $\delta_D \in [7, 28]$. However, when improving the detection system of infected individuals, these days may change since these individuals may be detected earlier. Hence, following [13], the functions are estimated as

$$\gamma_I(t) = \frac{1}{d_I - d_g(1-m_I(t))}, \quad \gamma_{I_u}(t) = \gamma_{H_R}(t) = \frac{1}{d_{H_R} + d_g(1-m_I(t))}, \quad \gamma_{H_D} = \frac{1}{d_{H_R}+\delta_D + d_g(1-m_I(t))},$$

where d_g is the maximum number of days that d_I can be reduced (i.e., $d_g = 6$). On the other hand, $\gamma_E = 0.1818 \approx 1/d_E$ remains constant. After many numerical tests, we set $\delta_D = 7$.

- **Initial condition:** During the study of the behaviour of the numerical results, we observed that, if we supposed a low number of infected people as initial condition, trying to suit the reported data, our numerical cases started to grow over 10 days later than the known data. Therefore, we needed to estimate some initial cases, although they do not correspond to what has been reported. Hence, we calibrated

$$S_0 = N - 5000,\ E_0 = 2480,\ I_0 = 1960,\ I_{u,0} = 360,\ H_{R,0} = 200,\ H_{D,0} = 0,\ R_{d,0} = R_{u,0} = D_0 = 0.$$

3.4.3 Numerical results

In this last Section, we present the numerical results obtained for the model using the numerical integrator Runge-Kutta 4 with time-step $dt = 0.01$ days, performing a trial-and-error method by hand.

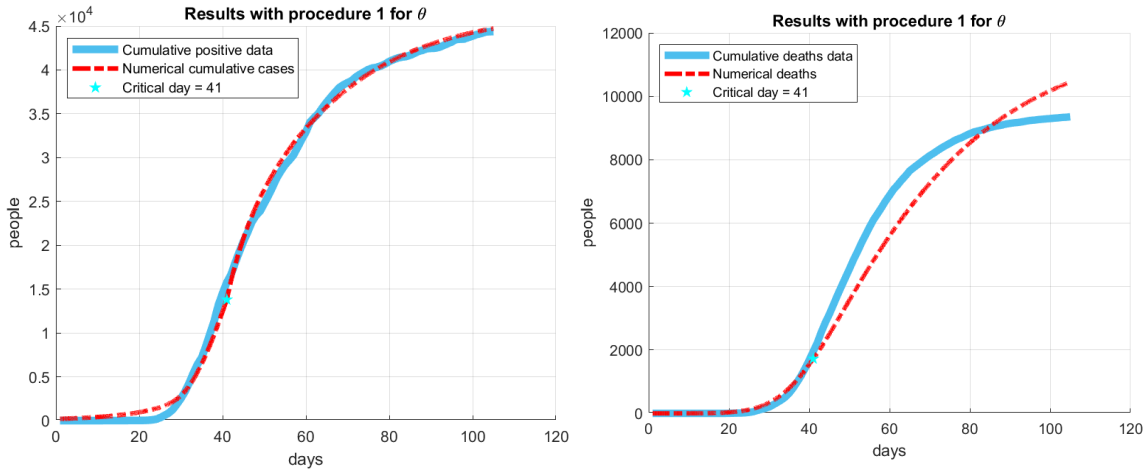


Figure 12: Numerical approximation of the reported data of positive cumulative cases and deaths during the pandemic of the coronavirus in the AMB using procedure 1.

In Figure 12, we show the results obtained using the calibrated parameters and the interpolation of θ and ω based on the prevalence (**procedure 1**). The integration has been performed from February 17^{th}, ten days before the first report of a positive case, until May 31^{st}, one day before the AMB entered Phase One. We observe that the cumulative cases are well approximated, but there is some error in the deaths cases; however, this could be due to an underreport of the *exitus* data.

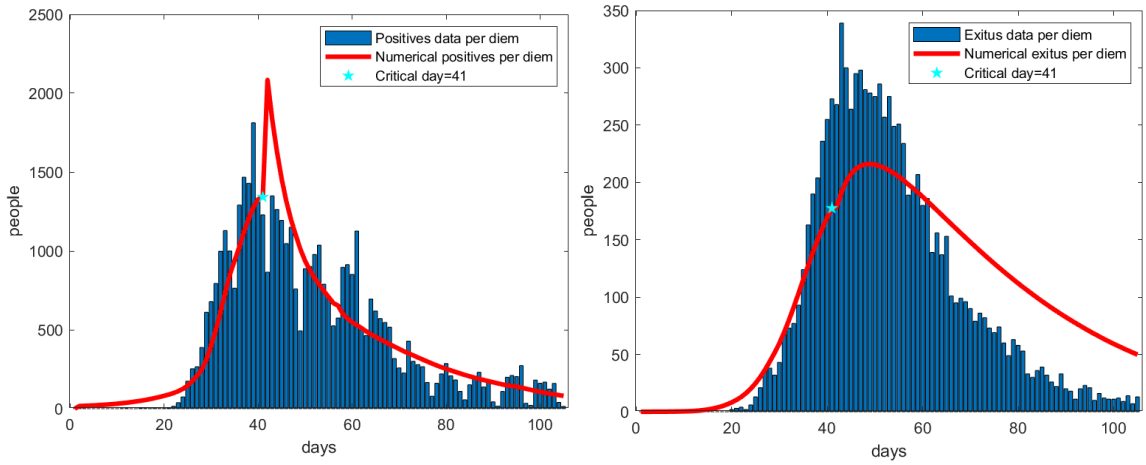


Figure 13: Daily reported cases *versus* numerical daily cases, for both positive cases and deaths.

Besides, we have marked the day corresponding to the effective reproduction number equal to 1, which means that the disease should start to refer; in fact, in Figure 12, this day corresponds (more or less) to the change of convexity of our numerical curves!

In Figure 13, we present the comparison between the daily reported cases and the ones obtained numerically, and it is observed that there is a delay of one day for the positive cases and around one week for the deaths to coincide with this critical $R_t = 1$, but it is still informative.

In Figure 14, we can see the evolution of the effective reproduction number R_t based on the

formula (11), with the aforementioned critical day when $R_t = 1$. This critical day $t = 41$ corresponds to March 28^{th}, from which the disease is supposed to start referring.

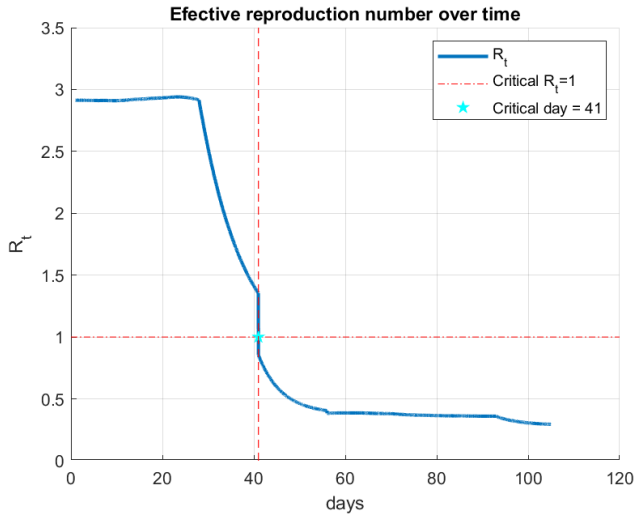


Figure 14: Numerical result for the effective reproduction number R_t based on equation (11) and procedure 1.

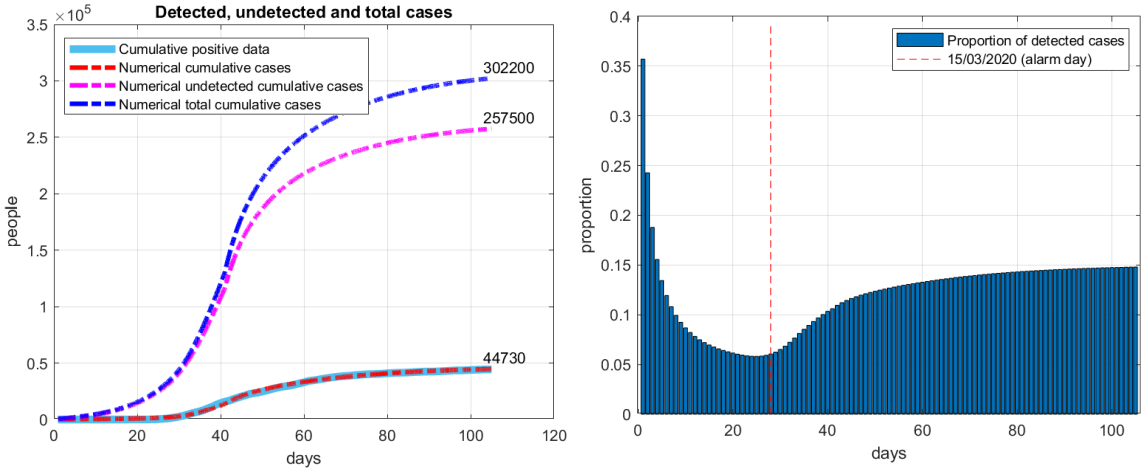


Figure 15: On the left, we present three curves: one for the estimated positive detected cases, another for the estimated undetected cases and a third one for the total estimated cases. On the right, the proportion of estimated detected cases.

Followingly, in Figure 15, we present the results of the uncertainty: our model estimates that, after the alarm day, the proportion of detected cases starts to stabilize on a 15% of the estimated total cases, over a 10% less than what estimated for θ in Figure 8. It is worth to notice that this stabilization starts around the alarm day. As aforementioned, these results must be interpreted carefully and keeping some doubt about them.

On the other hand, in Figure 16, we present the numerical results obtained by using the linear estimation for θ shown in Figure 9 and the smooth interpolation for ω shown in Figure 11 (**procedure 2**). Besides, all the other parameters are maintained the same. It is observed that the

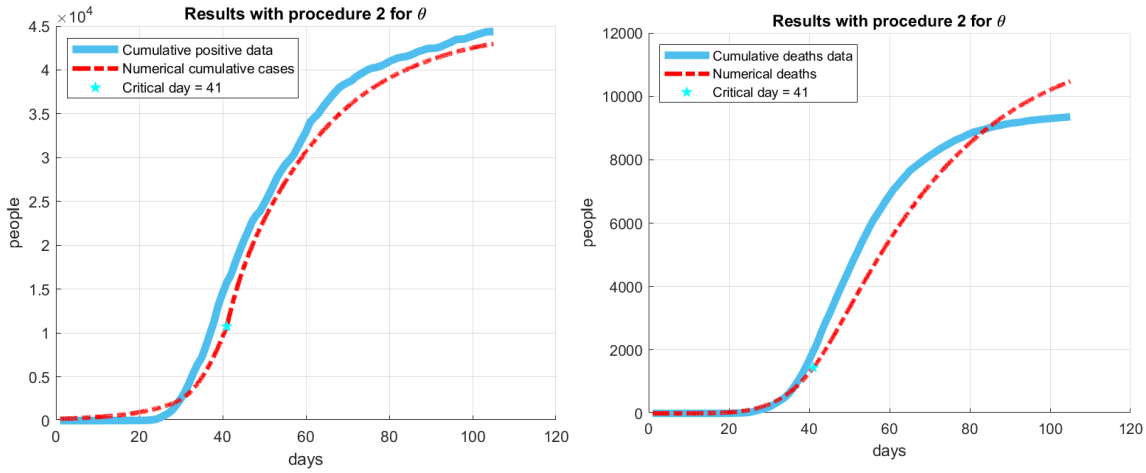


Figure 16: Numerical approximation of the reported data of positive cumulative cases and deaths during the pandemic of the coronavirus in the AMB using procedure 2.

results are very similar to the ones obtained in Figure 12, although some more calibration could be done in order to finish fitting them - in particular, the estimation of the positive cases is underrated. Nevertheless, the goal of these results was testing the accuracy of the second estimations for θ and ω in order to know if they are a good approach and we can make further *predictions* based on them.

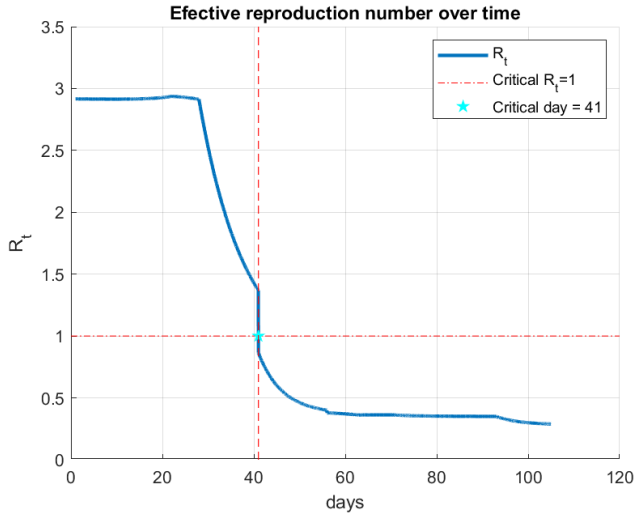


Figure 17: Numerical result for the effective reproduction number R_t based on equation (11) and procedure 2.

Furthermore, in Figure 17 it is presented the evolution of the effective reproduction number related to these results and it is observed that almost no change is appreciated with respect to Figure 14.

In Figure 18, we present the results for the second procedure after having re-calibrated some of the parameters in Section 3.4.2: in this case, $m_{I,1} = 0.33$, $m_{I,2} = 0.1$, $m_{I,3} = 0.105$, $m_{I,5} = 0.1$ and $m_{E,i} = m_{I,i} + 0.05$, $i \in \{1, \ldots, 5\}$. This re-calibration has improved the approximation for the

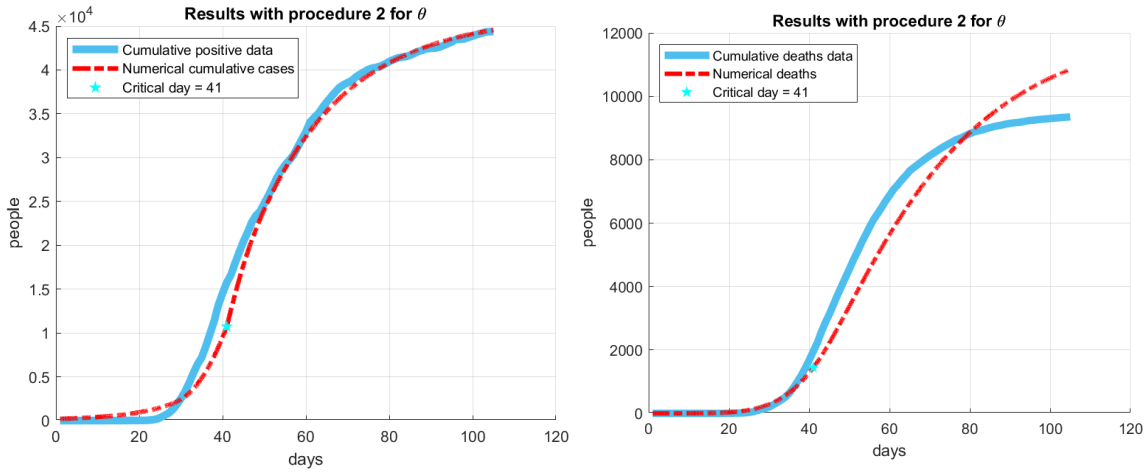


Figure 18: Numerical approximation of the reported data of positive cumulative cases and deaths during the pandemic of the coronavirus in the AMB using procedure 2 and re-calibrating some of the parameters.

positive cumulative cases in comparison with the results in Figure 16.

Let us now comment on the importance of the control measures by performing two different tests:

- **Late imposition of the control measures:** In Figure 19, we present the numerical results obtained when imposing the first control measures one week later (i.e., on March 22^{nd}). It is observed that our model's predictions are that imposing the control measures one week later could have supposed over $20,000$ more positive cases and $4,000$ more deaths with respect to the results in Figure 12.

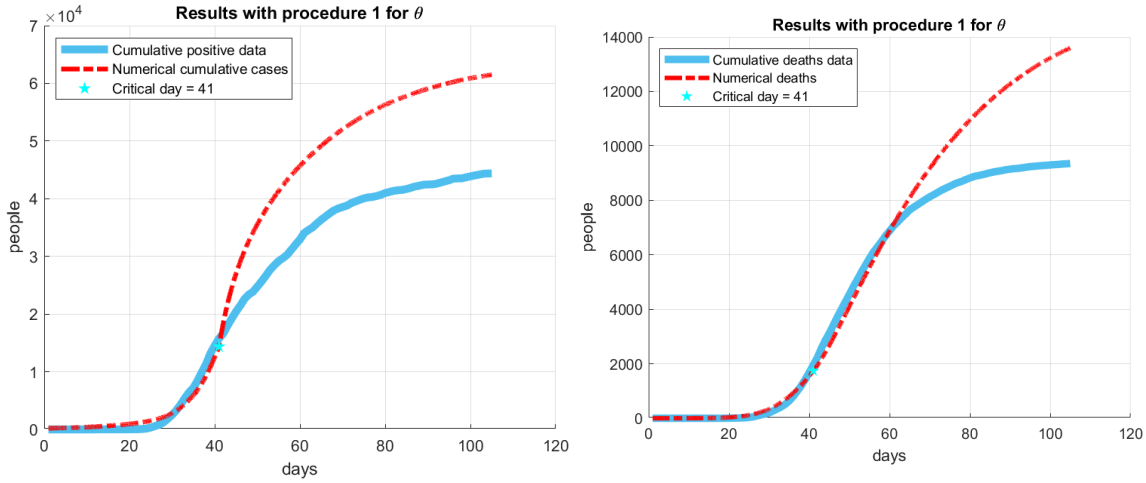


Figure 19: Numerical results using procedure 1 and supposing the control measures were applied one week later, on March 22^{nd}.

- **Rapid relaxation of the control measures:** In Figure 20, we present a *prediction* of the behaviour of the disease during 30 more days (i.e., until June 30^{th}). However, to perform this test, we have supposed a relaxation of the measure controls until a 40% less of effectiveness ($m_{I,6} = 0.48$ and $m_{E,6} = 0.52$). In Figure 21, we appreciate that this has led to a new increase on the effective reproduction number R_t, reaching $R_t = 1$ on the day 113, which corresponds to June 8^{th}. This corresponds to a regrowth of the disease in Figure 20. Notice that we have added the data until June 20^{th} and there is some correspondence until the regrowth.

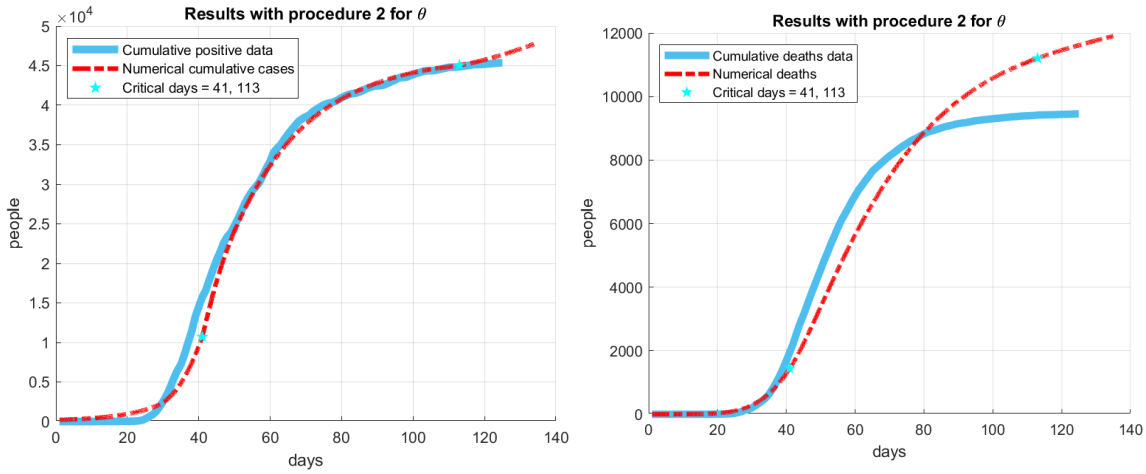


Figure 20: Numerical results obtained after considering a rapid relaxation of the control measures (a 40% less of effectiveness).

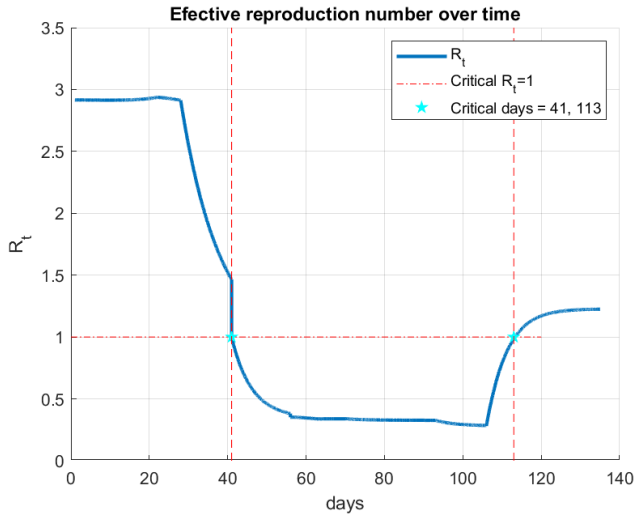


Figure 21: Numerical result for the effective reproduction number when relaxing rapidly the control measures for 30 days.

With these two tests concerning the control measures, we are trying to illustrate the importance of having appropriate protocols to fight the disease and having a collective social mind to attend the

authorities' commands in order to stop the spread. Furthermore, for our last test we have compared with the data until June 20^{th} and there was some accuracy at the beginning; however, we expect that, considering the real behaviour, the relaxation has not been so abrupt, which we will be able to check in the following days.

4 Conclusion and perspectives

In this work, we have presented some basic theory on mathematical epidemiological models and have tried to apply it on a θ-SEIHRD model adapted to the COVID-19. This model has some interesting characteristics, such as considering a compartment for undetected cases to save the uncertainty or the effect of the imposition of control measures. Besides, it works with a very relevant parameter, θ, corresponding to the proportion of detected cases; this allows us to have an intuition on the real magnitude of the pandemic.

Particularly, we have tried to fit the parameters of the model for the Metropolitan Area of Barcelona. The most remarkable result may be the estimation of a $75\% \sim 85\%$ of undetected cases in the area, although, as previously commented, these results are first intuitions and must be considered under some doubt. Besides, we have illustrated the importance of having appropriate protocols to minimize the damage, this is, imposing effective control measures and being careful when starting to relax them. Concretely, we have showed that (1) if we had applied these control measures one week later, there could have been over $\sim 20,000$ more (detected) positive cases and over $\sim 4,000$ more deaths, and (2) it is important to relax patiently these measures in order to avoid another regrowth in the following days.

All this study presented on the COVID-19 is just the beginning of a long period of studies, tests and predictions. It is worth to remark the importance of having consistent ways of reporting the data for the scientists to be able to present some intuition on the further evolution of this disease - the results here presented are attached to the data reported by the *Departament de Salut* and may not be adjusted for other different protocols (for instance, the *Ministerio de Sanidad* of the Spanish Government has been reporting different data for Catalunya with which θ and ω would have been probably lower).

Concerning on what has been performed in this work, there is still a long way to go to keep on performing these models; for instance:

- It is important to optimize the fitting algorithms in order to be able to work with and calibrate so many parameters; the calibration in this work has been done by hand by trial-and-error due to the computational cost that supposed an automatization, and the results may be very improvable in a near future with the development of these algorithms.
- The recent end of the alarm state leads to the need of considering outgoing and incoming flux in the cities to have a better control of the spread of the disease, which will add a greater degree of uncertainty.

It is necessary that each component on the society (politicians, scientists, citizens, etc.) work together and construct a solid strategy if we want to stop the spread and get rid of this pandemic.

References

[1] About the first epidemics in History: http://www.sld.cu/galerias/pdf/sitios/bmn/las_primeras_epidemias_de_la_historia. pdf, Red de Salud de Cuba.

[2] F. Brauer, *Mathematical epidemiology: Past, present and future*, Infectious Disease Modelling (2017) [Vol. 2, pp. 113-127], University of British Columbia, Vancouver, BC, Canada. https://doi.org/10.1016/j.idm.2017.02.001.

[3] The first known quarantined city in the History: https://www.lavanguardia.com/historiayvida/20200201/473229638796/cuarentena- coronavirus-gripe-contagio.html (April 12^{th} 2020), La Vanguardia, Barcelona.

[4] W. O. Kermack & A. G. McKendrick, *A contribution to the mathematical theory of epidemics*, Royal Society (August 1^{st} 1927) [Vol. 115, No. 772, pp. 700-721], Scotland. Republished in the *Bulletin of Mathematical Biology* (1991). https://doi.org/10.1098/rspa.1927.0118.

[5] F. Galton & H. W. Watson, *On the probability of the Extinction of Families*, The Journal of the Anthropological Institute of Great Britain and (1875) Ireland [Vol. 4, pp. 138-144]. https://doi.org/10.2307/2841222.

[6] P. van den Driessche & J. Watmough, *Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission* (2002), Mathematical Biosciences, Elsevier, Canada. https://doi.org/10.1016/S0025-5564(02)00108-6.

[7] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos* (1990) [Ch. 3, p. 37], Springer, Berlin.

[8] O. Diekmann, J. A. P. Heesterbeek & J. A. J. Metz, *On the Definition and the Computation of the Basic Reproduction Ratio R_0 in Models For Infectious-Diseases in Heterogeneous Populations*, Journal of Mathematical Biology (June 1990) [Vol. 28, 365-382], The Netherlands. https://doi.org/10.1007/BF00178324.

[9] P. van der Driessche & Z. Shuai, *Global stability of infectious disease models using Lyapunov functions*, SIAM Journal on Applied Mathematics (2013) [Vol. 73, No. 4, pp. 1513-1532], Canada. https://doi.org/10.1137/120876642.

[10] J. P. LaSalle, *The stability of Dynamical Systems* (1976) [Ch. 4, p. 46], Brown University, Providence. https://doi.org/10.1137/1.9781611970432.

[11] P. Hartman, *Ordinary Differential Equations* (1964) [Ch. 7, pp. 244-251], The Johns Hopkins University, Baltimore. https://doi.org/10.1137/1.9780898719222.

[12] A. Lyapunov, *Obshchaya zadacha ob ustoychivosti dvizheniya* (1892), University of Moscow, Moscow. Translation: A. T. Fuller, *The general problem of the stability of motion*, International Journal of Control (1992) [Vol. 55, No. 3, pp. 531-773]. https://doi.org/10.1080/00207179208934253.

[13] M. R. Ferrández, B. P. Ivorra, A. M. Ramos & M. Vela-Pérez, *Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) taking*

into account the undetected infections. The case of China, Communications in Nonlinear Science and Numerical Simulation (April 2020) [Vol. 88, 105303], Elsevier, Spain, https://doi.org/10.1016/j.cnsns.2020.105303.

[14] World Health Organization Director-general's opening remarks at the media briefing on COVID-19: https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks- at-the-media-briefing-on-covid-19--11-march-2020 (March 11^{th} 2020).

[15] Declaration of the alarm state in Spain: https://www.boe.es/diario_boe/txt.php?id=BOE-A-2020-3692 (March 14^{th} 2020), Madrid.

[16] Crucial dates of the quarantine in Spain: https://es.wikipedia.org/wiki/Cuarentena_de_Espa%C3%B1a_de_2020.

[17] Results of the first seroprevalence test in Spain by the *Instituto de Salud Carlos III*: https://www.mscbs.gob.es/ciudadanos/ene-covid/docs/ESTUDIO_ENE-COVID19_PRIMERA_RONDA_INFORME_PRELIMINAR.pdf (May 13^{th} 2020).

[18] On the compulsory use of masks in Spain: https://www.boe.es/buscar/doc.php?id=BOE-A-2020-5142 (May 19^{th} 2020), Madrid.

[19] Data source for the AMB: https://salutweb.gencat.cat/ca/inici, Departament de Salut, Generalitat de Catalunya, Barcelona.

[20] On the rests of SARS-CoV-19 found in Catalan waters: https://www.lavanguardia.com/local/barcelona/20200607/481654145204/detectados-restos-coronavirus-covid19-aguas-residuales-barcelona-enero.html, (June 8^{th} 2020), La Vanguardia, Barcelona.

[21] M. R. Ferrández, B. P. Ivorra, A. M. Ramos & M. Vela-Pérez, *Estimación de la evolución de la propagación de COVID-19 en la Comunidad de Madrid* (April 2020) [preprint], Spain. https://doi.org/10.13140/RG.2.2.27591.44969.

[22] Fernando Simón on the new corrections of the data in Spain: https://www.lavanguardia.com/vida/20200526/481420108410/fernando-simon-coronavirus-conteo-fallecidos-contagios.html (May 26^{th} 2020), La Vanguardia.

[23] Report about the Spanish healthcare workers by the *Instituto de Salud Carlos III*: https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/Documents/INFORMES/Informes%20COVID-19/COVID-19%20en%20Espa%C3%B1a.%20Situaci%C3%B3n%20en%20Sanitarios%20a%2021%20de%20mayo%20de%202020.pdf (May 21^{st} 2020).

[24] Epidemiological information for the COVID-19: https://www.ecdc.europa.eu/en/covid-19/latest-evidence/epidemiology, European Centre for Disease Prevention and Control.