BACHELOR'S THESIS

Degree in Mathematics

Title: Study of epidemic waves in a SEIR model

Author: Teresa Feu Basilio

Advisors: Gemma Huguet Casades and Jose Tomás Lázaro Ochoa

Department: Mathematics

Academic year: 2020-2021

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Facultat de Matemàtiques i Estadística

Universitat Politècnica de Catalunya
Facultat de Matemàtiques i Estadística

Degree in Mathematics

Bachelor's Degree Thesis

Study Of Epidemic Waves In A Seir Model

Teresa Feu Basilio

Supervised by Gemma Huguet Casades and Jose Tomás Lázaro Ochoa

June, 2021

Moltes gràcies als meus tutors Gemma i Tomás pel seu constant suport, ajuda i motivació durant aquest projecte. Gracias a mi madre, mi padre, y mis hermanas Silvia y Ana por su apoyo incondicional e increíble paciencia. Gracias a mis amigos y compañeros de la facultad por acompañarme durante todo este camino.

A mi Yaya, que se fue surfeando la primera ola.

Abstract

Many infectious diseases like the Spanish flu (1918-1919) or currently SARS-CoV-2, also known as COVID19, exhibit a certain wave-like behaviour. An epidemic wave can be defined as the time-distance between two consecutive peaks of infectious population. In many cases there is a seasonal reason for a wave-like behaviour. In others, like in COVID19, this is not so clear and it could be related to several circumstances: mobility restrictions imposed by governments, general lockdown, social contact constraints, school year, holidays... The goal of this Bachelor's degree thesis is to study epidemic waves from a dynamical systems approach, considering a simple SEIR model. We start by exploring theoretical results associated to a simple SEIR model as the basic reproduction number, Hethcote's theorem or variational equations. Then, we study how the SEIR parameter infection rate can lead to epidemic waves. Finally, we propose an original numerical method to apply our simple SEIR model to Catalonia's data. This method helps us analyse the importance of restrictions and to examine how small variations in the infection rate can lead to higher or smaller epidemic waves.

Keywords

Ordinary Differential Equations, Dynamical systems, SEIR model, Epidemiology, COVID19

Contents

1. Introduction

Epidemic waves caused by a disease are not a novelty. Among others, we can name the Spanish flu (1918-1919) which caused more than 500 million infected people, approximately one third of the world's population [11]. This exceptionally severe disease was responsible for an estimated number of 50 million deaths. Measures like social distancing or the use of face masks were imposed in order to reduce the impact of this influenza pandemic along its three waves of infectious cases. The first epidemic wave arose in the spring of 1918 and was rapidly followed by larger, higher and much more fatal second and third waves in the fall of 1918 and winter of 1919, respectively.

One hundred years later, another epidemic hit and stopped the world. This time it was the turn of SARS-COV2, also named COVID19. First detected in December 2019, COVID19 has caused more than 177 million infections and is responsible for almost 4 million deaths [12]. All around the world epidemic waves were detected when analysing the number of infectious cases over time. In particular, Spain has been among the most severelly hit countries in the world while experiencing three waves of infectious cases. Oddly enough, these three waves arose at the same period of the year as the Spanish flu waves. Indeed, the first wave took place from March 2020 until May 2020; the second from October 2020 until November 2020 and the third one from December 2020 until February 2021. In fact, when this study was proposed, Spain's second wave was just about to end and the dynamics of COVID19 epidemic waves triggered curiosity. Unlike other infectious diseases like influenza, there seemed not to be a seasonal reason for COVID19's wave-like behaviour. Hence, it could be related to several circumstances such as mobility restrictions imposed by governments, general lockdown, social distancing, school year, holidays...

Widely used in epidemiology, the SEIR model appeared to be the best fit to study COVID19 epidemic waves. In fact, this model is currently used by many scientists such as the ones from CSIC-UV [1] in order to model COVID19. However, in these models, parameters tend to be added or changed from the classic SEIR so that different phases such as hospitalizations or vaccinations can be considered. Therefore, we wondered to which extent a classic SEIR could model epidemic waves. In addition, choosing a simple SEIR would make us benefit from numerous theoretical theorems and properties.

In this final degree thesis we will study the dynamics of epidemic waves in a SEIR model by first exploring theoretical results and then applying our model to Catalonia's COVID19 data.

2. SEIR Model

The SEIR Model is a classic epidemiology model used to forecast the evolution of infectious diseases. This model belongs to the family of compartmental models. First introduced by Kermack and McKendrick in 1927[8], the compartmental models rely on the rates of flow between different groups of the population. Among them, we can find the models SIR, SIS, SEIS and SEIR. Indeed, the SEIR model consists in dividing the population in groups based on the condition of the individuals:

- Susceptible. The susceptible (S) is the part of the population that could be potentially subjected to the infection and has not yet been exposed to it.
- Exposed. The exposed (E) are individuals that have been exposed to the disease but are not yet infectious. They carry a very small number of pathogens and therefore are unable to transmit the disease. This period of time is called the latent phase.
- Infectious. The infectious (I) are individuals that have been infected with the disease and can actively transmit it.
- Recovered/Deceased. The recovered (R) represent individuals that have overcome the disease and whose probability of getting reinfected is very low. On the other hand, the deceased (D) is the fraction of infectious people who died from the disease.

The SEIR model lies on the transitions S $\longrightarrow$ E, E $\longrightarrow$ I and I $\longrightarrow$ R. It is a clear example of the so-called compartmental models.

The step S $\longrightarrow$ E involves the transmission of the disease which is determined by three factors: the prevalence[1] of the infected, the structure of contact of the population and the probability of transmission given a contact. These factors are represented by the parameter β, also called the *infection rate*. We can define β as $\beta = pb$ where p is the probability of transmission of the disease given a contact and b is the contact rate (the number of people, in average, in contact with an infectious individual) [6]. In other words, β is the number of people that an infectious person infects each day.

The next step, E $\longrightarrow$ I, lies on the latent phase. It is represented by the inverse of the average latent time γ which governs the lag between becoming infected and showing symptoms.

The final step is I $\longrightarrow$ R, where we include in R the deceased D. This step is governed by the recovery rate λ and the death rate κ. In the classic SEIR model they are considered as a unique parameter μ.

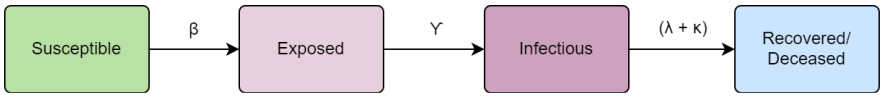


Figure 1: Classic SEIR model scheme

[1] Prevalence is the proportion of individuals in a population having a disease or specific characteristic in a given period of time.

2.1 SEIR Model equations

The equations for the classic SEIR model explained above are:

$$\begin{cases} \dfrac{dS(t)}{dt} = -\beta I \dfrac{S}{N} \\ \dfrac{dE(t)}{dt} = \beta I \dfrac{S}{N} - \gamma E \\ \dfrac{dI(t)}{dt} = \gamma E - (\lambda + \kappa) I \\ \dfrac{dR(t)}{dt} = (\lambda + \kappa) I \end{cases} \tag{1}$$

where N is the total population and $N = S + E + I + R$. We consider $\dot{N} = 0 \implies N = constant$. In other words, the total variation of the population can be underestimated in reference to the population size and, hence, we can assume that for every death there is a newborn. For a normalised SEIR model (i.e. $N = 1$), the equations would end up being:

$$\begin{cases} \dfrac{dS(t)}{dt} = -\beta I S \\ \dfrac{dE(t)}{dt} = \beta I S - \gamma E \\ \dfrac{dI(t)}{dt} = \gamma E - (\lambda + \kappa) I \\ \dfrac{dR(t)}{dt} = (\lambda + \kappa) I \end{cases} \tag{2}$$

Remark 2.1. To ease the notation during this study, sometimes we define $\mu = \lambda + \kappa$.

In the previous equations, we considered the recovered (R) and the deceased (D) as a same group in our dynamics. However, it is possible to differentiate these two groups to expand the study. The scheme in figure 2 below illustrates this situation.

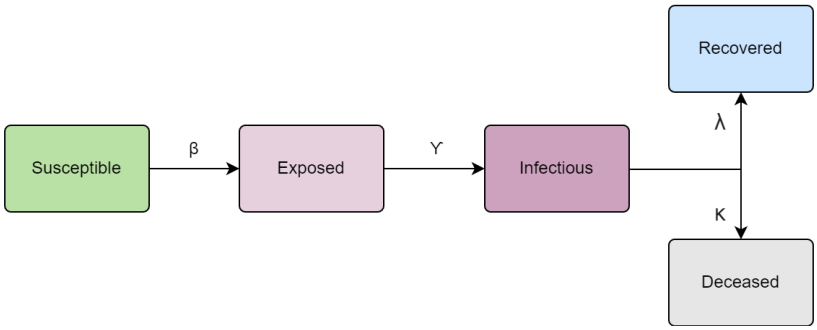


Figure 2: SEIR model differentiating (R) from (D) scheme

Hence, when differentiating the recovered from the deceased, the following equations are derived:

$$\begin{cases} \frac{dS(t)}{dt} = -\beta I \frac{S}{N} \\ \frac{dE(t)}{dt} = \beta I \frac{S}{N} - \gamma E \\ \frac{dI(t)}{dt} = \gamma E - (\lambda + \kappa) I \\ \frac{dR(t)}{dt} = \lambda I \\ \frac{dD(t)}{dt} = \kappa I \end{cases} \tag{3}$$

where $N = S + E + I + R = constant$.

All this time, we have assumed that the total population $N(t)$ was a constant. In other words, we have supposed that for every death there was a newborn. However, there could be many deaths due to COVID19 not offset by the number of newborns. In that case, we would want $N(t)$ to be a variable where $N(t) = S(t) + E(t) + I(t) + R(t)$. This would lead us to the equations:

$$\begin{cases} \frac{dS(t)}{dt} = -\beta I \frac{S}{N(t)} \\ \frac{dE(t)}{dt} = \beta I \frac{S}{N(t)} - \gamma E \\ \frac{dI(t)}{dt} = \gamma E - (\lambda + \kappa) I \\ \frac{dR(t)}{dt} = \lambda I \\ \frac{dD(t)}{dt} = \kappa I \end{cases} \tag{4}$$

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

In our study, we considered working with these three models. However, in order to be able to apply the different theorems and properties that we will present in the following sections, we decided to work exclusively with the classic SEIR Model equations (1) and (2).

2.2 Basic reproduction number R_0

The basic reproduction number is of high relevance in epidemiology since from its value we can know whether the disease cannot grow anymore or it can quickly spread through the population.

Definition 2.2. The basic reproduction number R_0 is the expected number of secondary cases produced by a typical infectious individual in a completely susceptible population.

In other words, if $R_0 \geq 1$, on average an infectious individual causes more than one new infected individual over the course of its infectious period. Hence, the disease will quickly spread through the population. If $R_0 < 1$ it means that an infectious individual will produce less than one new infected individual and the disease can be controlled.

Definition 2.3. A disease free equilibrium (DFE) is an equilibrium point x^* where there is no disease.

For instance, in the classic SEIR model a DFE x^* is such that $x^* \subset \{x = (S, E, I, R) | E = 0, I = 0\}$.

Definition 2.4. The basic reproduction number can also be defined as the number of new infections produced by a typical infectious individual in a population around a DFE. [14]

Our aim in this section is to find an expression for the basic reproduction number R_0 following definition 2.4. We will explore two different approaches in order to reach our goal.

2.2.1 Classic method

Let us recall the following theorem:

Theorem 2.5. *Let us consider the non linear system $\dot{x} = f(x)$. Let x^* be an equilibrium point of this system and let A be the matrix*

$$A = Df(x^*) = \left(\frac{\partial f_i}{\partial x_j}(x^*)\right)_{1 \leq i,j \leq n}$$

1. *If A has an eigenvalue λ with $\mathcal{R}e(\lambda) > 0$, then x^* is unstable.*
2. *If every eigenvalue λ of A has $\mathcal{R}e(\lambda) < 0$, then x^* is asymptotically stable.*

Following this theorem, we proceed to study the stability of a disease free equilibrium point of the SEIR system. We start by calculating the Jacobian matrix J of the classic SEIR system:

$$J = \begin{pmatrix} \frac{-\beta I}{N} & 0 & \frac{-\beta S}{N} & 0 \\ \frac{\beta I}{N} & -\gamma & \frac{\beta S}{N} & 0 \\ 0 & \gamma & -(\lambda + \kappa) & 0 \\ 0 & 0 & \lambda + \kappa & 0 \end{pmatrix}$$

We will consider the equilibrium point $x^* = (S, 0, 0, N - S)$ of the invariant manifold $\{E = 0, I = 0\}$. The equilibrium point x^* corresponds to the situation where there are no exposed or infectious cases. Our population is either susceptible or recovered and we have overcome the virus.

Let J_{x^*} be the value of the Jacobian matrix J at the equilibrium point x^*. Note that $J_{x^*} = A$.

$$J_{x^*} = \begin{pmatrix} 0 & 0 & \frac{-\beta S}{N} & 0 \\ 0 & -\gamma & \frac{\beta S}{N} & 0 \\ 0 & \gamma & -(\lambda+\kappa) & 0 \\ 0 & 0 & \lambda+\kappa & 0 \end{pmatrix}$$

We calculate the eigenvalues c_i of J_{x^*} by solving the equation $det(J_{x^*} - c \cdot \mathrm{Id}) = 0$.

$$\begin{aligned} det(J_{x^*} - cId) = 0 &\iff \begin{vmatrix} -c & 0 & \frac{-\beta S}{N} & 0 \\ 0 & -\gamma - c & \frac{\beta S}{N} & 0 \\ 0 & \gamma & -(\lambda+\kappa) - c & 0 \\ 0 & 0 & \lambda+\kappa & -c \end{vmatrix} = 0 \\ &\iff (-c) \begin{vmatrix} -\gamma - c & \frac{\beta S}{N} & 0 \\ \gamma & -(\lambda+\kappa) - c & 0 \\ 0 & \lambda+\kappa & -c \end{vmatrix} = 0 \\ &\iff c^2 \begin{vmatrix} -\gamma - c & \frac{\beta S}{N} \\ \gamma & -(\lambda+\kappa) - c \end{vmatrix} = 0 \\ &\iff c^2 \left(c^2 + c(\lambda+\kappa+\gamma) + \gamma(\lambda+\kappa) - \gamma\frac{\beta S}{N} \right) = 0 \\ &\iff \begin{cases} c = 0 \\ c^2 + c(\lambda+\kappa+\gamma) + \gamma(\lambda+\kappa) - \gamma\frac{\beta S}{N} = 0 \end{cases} \\ &\iff \begin{cases} c = 0 \\ c = \dfrac{-(\lambda+\kappa+\gamma) \pm \sqrt{(\lambda+\kappa+\gamma)^2 - 4\gamma(\lambda+\kappa-\frac{\beta S}{N})}}{2} := c_\pm \end{cases} \end{aligned}$$

Hence, J_{x^*} has four eigenvalues: $c_1 = c_2 = 0$, $c_3 = c_-$ and $c_4 = c_+$.

We need to distinguish two situations:

1. if $(\lambda+\kappa+\gamma)^2 < 4\gamma(\lambda+\kappa-\frac{\beta S}{N})$, then $c_\pm \in \mathbb{C}$ and

   $$c_\pm = \frac{-(\lambda+\kappa+\gamma) \pm i\sqrt{4\gamma(\lambda+\kappa-\frac{\beta S}{N}) - (\lambda+\kappa+\gamma)^2}}{2}$$

   We observe that in this case $\mathcal{R}e(c_\pm) < 0$.

2. if $(\lambda+\kappa+\gamma)^2 \geq 4\gamma(\lambda+\kappa-\frac{\beta S}{N})$, then $c_\pm \in \mathbb{R}$. In this second case, $c_- < 0$. However, if $\lambda+\kappa-\frac{\beta S}{N} < 0$ we observe that $c_+ > 0$. Moreover,

   $$c_+ > 0 \iff \lambda+\kappa < \frac{\beta S}{N} \iff \frac{\beta S}{N(\lambda+\kappa)} > 1,$$

   where we have taken into account that $\lambda+\kappa > 0$. Therefore, if $\dfrac{\beta S}{N(\lambda+\kappa)} > 1$, x^* is unstable.

We can now define the basic reproduction number R_0 as

$$R_0 = \frac{\beta S}{N(\lambda+\kappa)}$$

2.2.2 Next-generation matrix

Even though the expression that we found in the previous section is correct, we omitted to study the stability of the eigenvalues equal to zero. However, the next-generation matrix (NGM), first introduced by Diekmann and Heesterbeek in [5], can be useful to solve this issue. Hence, in this section we will study this method following the articles written by Sallet [13], Van den Driessche and Watmough [14].

Let us suppose that we divide our population in several compartments where the m first compartments correspond to uninfected people and the last ones to infected people. Let $x = (x_1, ..., x_n)^T$, with each $x_i \geq 0$, be the number of individuals in each compartment. Let Y_s be the set of all disease free states, $Y_s = \{x \geq 0 \mid x_i = 0, i = m+1, ..., n\}$.

Let $\mathcal{F}_i(x)$ be the rate of appearance of new infections in compartment i, $\mathcal{V}_i^+(x)$ the rate of transfer of individuals into compartment i by all other means, and $\mathcal{V}_i^-(x)$ the velocity of transfer of individuals out of compartment i. Let $\mathcal{V}_i = \mathcal{V}_i^+ - \mathcal{V}_i^-$. We have $\dot{x}_i = f(x) = \mathcal{F}_i(x) + \mathcal{V}_i(x)$ for $i = 1, ..., n$. From the epidemiological nature of the model follow the properties below:

1. $x \geq 0, \mathcal{F}_i(x) \geq 0, \mathcal{V}_i^+(x) \geq 0, \mathcal{V}_i^-(x) \geq 0$
2. If $x_i = 0$ then $\mathcal{V}_i^-(x) = 0$
3. If $i \leq m$ then $F_i(x) = 0$
4. If $x \in Y_s$ then $F_i(x) = 0$ and for $i \geq m$ we have $\mathcal{V}_i^+(x) = 0$, $\mathcal{V}_i^-(x) = 0$ and $\mathcal{F}_i(x) = 0$

Let X be the array of the infected compartments and Y be the array of all disease free compartments such as $x = (Y, X)^T$. Our system can be rewritten as

$$\begin{cases} \dot{X} = f_1(X, Y) \\ \dot{Y} = f_2(X, Y) \end{cases}$$

In our case, we have

$$X = \begin{pmatrix} E \\ I \end{pmatrix}, Y = \begin{pmatrix} S \\ R \end{pmatrix}$$

from which follows

$$\mathcal{F} = \begin{pmatrix} 0 \\ 0 \\ \beta \frac{SI}{N} \\ 0 \end{pmatrix}, \mathcal{V} = \begin{pmatrix} -\beta \frac{SI}{N} \\ (\lambda + \kappa) I \\ -\gamma E \\ \gamma E - (\lambda + \kappa) I \end{pmatrix}$$

Observation 2.6. A DFE is $x^* \in Y_s$ such that $f(x^*) = 0$.

We consider now the linearization of $\dot{x} = f(x)$ around a DFE x^*. In our case we have that any $x^* = (S_0, N - S_0, 0, 0)$ is a DFE.

$$Df(x) = \begin{pmatrix} -\frac{\beta I}{N} & 0 & 0 & -\frac{\beta S}{N} \\ 0 & 0 & 0 & \lambda + \kappa \\ \frac{\beta I}{N} & 0 & -\gamma & \frac{\beta S}{N} \\ 0 & 0 & \gamma & -(\lambda + \kappa) \end{pmatrix} \Longrightarrow Df(x^*) = \begin{pmatrix} 0 & 0 & 0 & -\frac{\beta S_0}{N} \\ 0 & 0 & 0 & \lambda + \kappa \\ 0 & 0 & -\gamma & \frac{\beta S_0}{N} \\ 0 & 0 & \gamma & -(\lambda + \kappa) \end{pmatrix}$$

Observation 2.7. Van den Driessche and Watmough [14] suppose that the DFE is locally asymptotically stable. However, this is not exactly true in our case. Therefore, we follow Sallet's reasoning [13].

Definition 2.8. Let F and V be matrices satisfying

$$D\mathcal{F}(x^*) = \begin{bmatrix} 0 & 0 \\ 0 & F \end{bmatrix}, D\mathcal{V}(x^*) = \begin{bmatrix} J_3 & J_4 \\ 0 & V \end{bmatrix}$$

In our case we have

$$D\mathcal{F}(x) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{\beta I}{N} & 0 & 0 & \frac{\beta S}{N} \\ 0 & 0 & 0 & 0 \end{pmatrix} \Longrightarrow D\mathcal{F}(x^*) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\beta S_0}{N} \\ 0 & 0 & 0 & 0 \end{pmatrix} \Longrightarrow F = \begin{pmatrix} 0 & \frac{\beta S_0}{N} \\ 0 & 0 \end{pmatrix}$$

$$D\mathcal{V}(x) = \begin{pmatrix} -\beta\frac{I}{N} & 0 & 0 & -\beta\frac{S}{N} \\ 0 & 0 & 0 & \lambda+\kappa \\ 0 & 0 & -\gamma & 0 \\ 0 & 0 & \gamma & -(\lambda+\kappa) \end{pmatrix} \Longrightarrow D\mathcal{V}(x^*) = \begin{pmatrix} 0 & 0 & 0 & -\beta\frac{S_0}{N} \\ 0 & 0 & 0 & \lambda+\kappa \\ 0 & 0 & -\gamma & 0 \\ 0 & 0 & \gamma & -(\lambda+\kappa) \end{pmatrix}$$

$$\Longrightarrow V = \begin{pmatrix} -\gamma & 0 \\ \gamma & -(\lambda+\kappa) \end{pmatrix}$$

Observation 2.9. F and V are defined by

$$F = \left(\frac{\partial \mathcal{F}}{\partial X}\right)_{x^*}, V = \left(\frac{\partial \mathcal{V}}{\partial X}\right)_{x^*}$$

Definition 2.10. A Metzler matrix is a matrix in which all the off-diagonal components are positive (equal to or greater than zero), that is to say if $i \neq j$ then $a_{ij} \geq 0$.

Theorem 2.11. *The matrix F is a positive matrix, i.e $\forall i, j$ $f_{ij} \geq 0$; and V is a Metzler matrix.*

Definition 2.12. The spectral radius of a matrix A is the supremum among the absolute values of the eigenvalues of A, that is to say

$$\rho(A) = \max_{\lambda \in Sp(A)} \mid \lambda \mid$$

Definition 2.13. We define R_0 as $R_0 = \rho(-FV^{-1})$. We call $-FV^{-1}$ the next generation matrix of the model.

In our case, we have

$$-FV^{-1} = \frac{1}{\gamma(\lambda+\kappa)} \begin{pmatrix} \gamma\frac{\beta S}{N} & \gamma\frac{\beta S}{N} \\ 0 & 0 \end{pmatrix}$$

Thus, $R_0 = \rho(-FV^{-1}) = \dfrac{\beta S}{N(\lambda+\kappa)}$.

The following definition and theorem assure the accuracy of the definition of R_0.

Definition 2.14. Let A be a Metzler matrix. A regular decomposition of A is a decomposition of the form

$$A = F + V$$

where $F \geq 0$ and V is a Metzler matrix asymptotically stable (equivalent to V invertible).

Theorem 2.15. *(Varga) For all regular decomposition of a Metzler invertible matrix, the following assertions are equivalent*

- *A is asymptotically stable*
- $\rho(-FV^{-1}) < 1$

2.3 Hethcote's theorem

Curious about the characteristics of a SEIR wave, we encountered Hethcote's article [7] in which lied an interesting theorem. Among other facts, this theorem affirms that a classic SEIR model can experience either one wave or none. In addition, it remarks the relevance of the infection rate β in the dynamics of a SEIR wave. Indeed, as we will see in section 3, variations of β can lead to several epidemic waves in a SEIR model.

Theorem 2.16. *Let (S(t), E(t), I(t), R(t)) be a solution of the classic SEIR model (1) with initial conditions* $S(0) = S_0, E(0) = E_0, I(0) = I_0$ *and* $R(0) = R_0$. *Let* $\mu = \lambda + \kappa$ *and* $\sigma = \dfrac{\beta}{N\mu}$. *If* $\sigma S_0 \leq 1$, *then* $E(t)$ *and* $I(t)$ *decrease to zero as* $t \to \infty$. *If* $\sigma S_0 > 1$, *then* $E(t) + I(t)$ *first increases up a maximum with value*

$$E_{max} + I_{max} = E_0 + I_0 + S_0 - \frac{1}{\sigma} ln(\sigma S_0) - \frac{1}{\sigma} \tag{5}$$

and then decreases to zero as $t \to \infty$.

Demonstration 2.17. Let us demonstrate the formula of the maximum value (5). Recall the equations of the classic SEIR model (1):

$$\begin{cases} \dfrac{dS(t)}{dt} = -\beta I \dfrac{S}{N} & (i) \\ \dfrac{dE(t)}{dt} = \beta I \dfrac{S}{N} - \gamma E & (ii) \\ \dfrac{dI(t)}{dt} = \gamma E - \mu I & (iii) \\ \dfrac{dR(t)}{dt} = \mu I & (iv) \end{cases}$$

From (i) and (iv) we have $\dfrac{S'}{S} = -\beta \dfrac{I}{S} = -\beta \dfrac{R'}{\mu N}$.

Thus, if we integrate and take into account that $R(0) = 0$ we obtain

$$\int \frac{S'}{S} = \int -\beta \frac{R'}{\mu N} \Longrightarrow ln\left(\frac{S(t)}{S(0)}\right) = \frac{-\beta}{N\mu} R(t) \tag{6}$$

Let t^* be the time when $E(t) + I(t)$ has a maximum. We know that $E'(t^*) + I'(t^*) = 0$. Moreover,

$$\begin{aligned} E'(t) + I'(t) = 0 &\Longleftrightarrow I\left(\beta \frac{S}{N} - \mu\right) = 0 \\ &\overset{I \neq 0}{\Longrightarrow} S(t^*) = \frac{\mu N}{\beta} = \frac{1}{\sigma} \end{aligned} \tag{7}$$

Now, from (6) we have

$$R(t^*) = \frac{-N\mu}{\beta} ln\left(\frac{S(t^*)}{S(0)}\right) \overset{(7)}{=} -\frac{1}{\sigma} ln\left(\frac{1}{\sigma S(0)}\right) = \frac{1}{\sigma} ln[S(0)]$$

We know that $E(t) + I(t) = N - R(t) - S(t)$, where $N = constant = S(0) + E(0) + I(0)$, therefore

$$E_{max} + I_{max} = N - R(t^*) - S(t^*) = S(0) + E(0) + I(0) - \frac{1}{\sigma} ln[S(0)] - \frac{1}{\sigma}$$

2.4 Variational equations

Our aim in this section is to study the variational equations regarding parameters of our SEIR model in order to apply them in section 4.4. As we will see, variational equations regarding parameters can help study how small variations in a given infection rate can influence the number of infectious people over time.

We will start by recalling the First Variational Equation (9) regarding initial conditions. Then we will study variationals equations regarding parameters, to finally apply them to our SEIR model.

2.4.1 Variationals regarding initial conditions

Let us consider the initial value problem

$$\begin{cases} \dot{x} = f(t,x) \text{ where } f : \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n, f \in \mathcal{C}^r, r \geq 2 \\ x(t_0) = x_0 \end{cases} \tag{8}$$

and let $x = \varphi(t; t_0, x_0)$ be a solution of the initial value problem (8), where $\varphi \in \mathcal{C}^r$, $r \geq 2$.

We know that

$$\frac{d}{dt}\left(\varphi(t; t_0, x_0)\right) = f(t, \varphi(t; t_0, x_0))$$

Therefore, applying the chain rule and considering $f, \varphi \in \mathcal{C}^r$, $r \geq 2$, we get

$$\frac{d}{dt}(D_{x_0}\varphi(t; t_0, x_0)) = D_{x_0}\left(\frac{d}{dt}(\varphi(t; t_0, x_0)\right) = D_x f(t, \varphi(t; t_0, x_0)) D_{x_0}\varphi(t; t_0, x_0)$$

Moreover,

$$\varphi(t_0; t_0, x_0) = x_0 \Longrightarrow D_{x_0}\varphi(t; t_0, x_0) = Id$$

In other words, $D_{x_0}\varphi(t; t_0, x_0)$ is a solution of the so called First Variational Equation (along $\varphi(t; t_0, x_0)$):

$$\begin{cases} \dot{Z} = D_x f(t; \varphi(t; t_0, x_0)) Z \\ Z(t_0) = Id \end{cases} \tag{9}$$

which provides information about the sensitivity of the solutions of $\dot{x} = f(t,x)$ when one starts with initial conditions close to x_0 (for $t = t_0$).

2.4.2 Variationals regarding parameters

In order to study the dependency of an initial value problem regarding a parameter $\mu \in \mathcal{U} \subset \mathbb{R}^p$, we consider μ as a variable and apply the previous formula.

Let us consider the initial value problem

$$\begin{cases} \dot{x} = f(t,x,\mu) \text{ where } f : \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n, f \in \mathcal{C}^r, r \geq 2, \mu \in \mathcal{U} \subset \mathbb{R}^p, x \in \mathbb{R}^n \\ x(t_0, \mu_0) = x_0 \end{cases} \tag{10}$$

and let $x = \varphi(t; t_0, x_0, \mu_0)$ be a solution of the initial value problem (10), where $\varphi \in \mathcal{C}^r$, $r \geq 2$.

Let y be $y = \begin{pmatrix} x \\ \mu \end{pmatrix}$. Then $\dot{y} = \begin{pmatrix} \dot{x} \\ \dot{\mu} \end{pmatrix} = \begin{pmatrix} f(t,y) \\ 0 \end{pmatrix} =: F(t,y)$.

In this case the First Variational Equation (VE1) would be

$$\begin{cases} \dot{W}(t) = D_y F(t, \begin{pmatrix} \varphi(t; t_0, x_0, \mu_0) \\ \mu_0 \end{pmatrix}) W(t) \\ W(t_0) = Id \end{cases} \tag{11}$$

Moreover,

$$\begin{aligned} D_y F(t, \begin{pmatrix} \varphi(t; t_0, x_0, \mu_0) \\ \mu_0 \end{pmatrix}) &= D_y \begin{pmatrix} f(t, \varphi(t; t_0, x_0, \mu_0), \mu_0) \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} D_x f(t, \varphi(t; t_0, x_0, \mu_0), \mu_0) & \frac{\partial}{\partial \mu} f(t, \varphi(t; t_0, x_0, \mu_0), \mu_0) \\ 0 & 0 \end{pmatrix} \end{aligned}$$

Then we can write VE1 as

$$\begin{pmatrix} D_{x_0}\varphi(t; t_0, x_0, \mu_0) & \frac{\partial}{\partial \mu}\varphi(t; t_0, x_0, \mu_0) \\ 0 & Id_p \end{pmatrix}' = \begin{pmatrix} D_x f(t, \varphi(t; t_0, x_0, \mu_0), \mu_0) & \frac{\partial}{\partial \mu} f(t, \varphi(t; t_0, x_0, \mu_0), \mu_0) \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} D_{x_0}\varphi(t; t_0, x_0, \mu_0) & \frac{\partial}{\partial \mu}\varphi(t; t_0, x_0, \mu_0) \\ 0 & Id_p \end{pmatrix}$$

where the initial conditions are

$$\begin{cases} D_{x_0}\varphi(t_0; t_0, x_0, \mu_0) = Id \\ \frac{\partial}{\partial \mu}\varphi(t_0; t_0, x_0, \mu_0) = 0 \end{cases}$$

Thus, we have

$$\frac{d}{dt}(D_{x_0}\varphi(t; t_0, x_0, \mu_0)) = D_x f(t, \varphi(t; t_0, x_0, \mu_0), \mu_0) D_{x_0}\varphi(t; t_0, x_0, \mu_0) \tag{12}$$

$$\frac{d}{dt}\left(\frac{\partial}{\partial \mu}\varphi(t; t_0, x_0, \mu_0)\right) = D_x f(t, \varphi(t; t_0, x_0, \mu_0), \mu_0)\frac{\partial}{\partial \mu}\varphi(t; t_0, x_0, \mu_0) + \frac{\partial}{\partial \mu} f(t, \varphi(t; t_0, x_0, \mu_0), \mu_0) \tag{13}$$

Let us apply these equations to our system

$$\begin{cases} \frac{dS(t)}{dt} = -\beta I \frac{S}{N} \\ \frac{dE(t)}{dt} = \beta I \frac{S}{N} - \gamma E \\ \frac{dI(t)}{dt} = \gamma E - (\lambda + \kappa) I \\ \frac{dR(t)}{dt} = (\lambda + \kappa) I \end{cases} \iff \dot{x} = f(x, \mu) \text{ where } \begin{cases} x = (S, E, I, R) \\ \mu = (\beta, \gamma, \lambda, \kappa) \end{cases}$$

with $N = S + E + I + R$, $\dot{N}(t) = 0$.

We consider $t_0 = 0$ since our system is autonomous. Then, if $\varphi(t, x_0, \mu_0)$ is a solution of the initial value problem

$$\begin{cases} \dot{x} = f(x, \mu) \\ x(0, \mu_0) = x_0 \end{cases}$$

we get

$$D_x f(\varphi(t, x_0, \mu_0)) = \begin{pmatrix} \frac{-\beta I}{N} & 0 & \frac{-\beta S}{N} & 0 \\ \frac{\beta I}{N} & -\gamma & \frac{\beta S}{N} & 0 \\ 0 & \gamma & -(\lambda + \kappa) & 0 \\ 0 & 0 & \lambda + \kappa & 0 \end{pmatrix}$$

$$\frac{\partial}{\partial \mu} f(x, \mu) = \begin{pmatrix} \frac{-IS}{N} & 0 & 0 & 0 \\ \frac{IS}{N} & -E & 0 & 0 \\ 0 & E & -I & -I \\ 0 & 0 & I & I \end{pmatrix}$$

Let $B(t)$ be the vector

$$B(t) = \frac{\partial}{\partial \beta} \varphi(t, x_0, \mu_0)^T = \big(b_1(t), b_2(t), b_3(t), b_4(t)\big)^T = \left(\frac{\partial S}{\partial \beta}, \frac{\partial E}{\partial \beta}, \frac{\partial I}{\partial \beta}, \frac{\partial R}{\partial \beta}\right)^T$$

From (13) we observe that $B(t)$ satisfies the following differential equation not homogeneous:

$$\dot{B}(t) = \begin{pmatrix} \frac{-\beta I}{N} & 0 & \frac{-\beta S}{N} & 0 \\ \frac{\beta I}{N} & -\gamma & \frac{\beta S}{N} & 0 \\ 0 & \gamma & -(\lambda + \kappa) & 0 \\ 0 & 0 & \lambda + \kappa & 0 \end{pmatrix} B(t) + \frac{1}{N} \begin{pmatrix} -IS \\ IS \\ 0 \\ 0 \end{pmatrix} \tag{14}$$

where $\varphi(t, x_0, \mu_0) = \big(S(t), E(t), I(t), R(t)\big)$

That is to say,

$$\dot{b_1}(t) = -\frac{\beta}{N} I(t) b_1(t) - \frac{\beta}{N} S(t) b_3(t) - \frac{1}{N} I(t) S(t) \tag{15}$$

$$\dot{b_2}(t) = \frac{\beta}{N} I(t) b_1(t) - \gamma b_2(t) + \frac{\beta}{N} S(t) b_3(t) + \frac{1}{N} I(t) S(t) \tag{16}$$

$$\dot{b_3}(t) = \gamma b_2(t) - (\lambda + \kappa) b_3(t) \tag{17}$$

$$\dot{b_4}(t) = (\lambda + \kappa) b_3(t) \tag{18}$$

We observe that the equations (15)-(16)-(17) decouple from the last one (18).

Let us define $W(t) = \begin{pmatrix} b_1(t) \\ b_2(t) \\ b_3(t) \end{pmatrix}$. From (14) we obtain

$$\dot{W}(t) = \begin{pmatrix} \frac{-\beta}{N} I(t) & 0 & \frac{-\beta}{N} S(t) \\ \frac{\beta}{N} I(t) & -\gamma & \frac{\beta}{N} S(t) \\ 0 & \gamma & -(\lambda + \kappa) \end{pmatrix} W(t) + \frac{1}{N} \begin{pmatrix} -I(t) S(t) \\ I(t) S(t) \\ 0 \end{pmatrix} = M(t) W(t) + \frac{1}{N} C(t) \tag{19}$$

We observe that $tr(M(t)) = \frac{-\beta}{N} I(t) \leq 0$, $\forall t \geq 0$.

Remark 2.18. At $t = t_0$, $W(t_0) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$. Indeed, if we consider β as a variable and we calculate the variationals regarding initial conditions we have:

$$\begin{pmatrix} \frac{\partial S}{\partial S_0} & \frac{\partial S}{\partial E_0} & \frac{\partial S}{\partial I_0} & \frac{\partial S}{\partial \beta_0} \\ \frac{\partial E}{\partial S_0} & \frac{\partial E}{\partial E_0} & \frac{\partial E}{\partial I_0} & \frac{\partial E}{\partial \beta_0} \\ \frac{\partial I}{\partial S_0} & \frac{\partial I}{\partial E_0} & \frac{\partial I}{\partial I_0} & \frac{\partial I}{\partial \beta_0} \\ \frac{\partial \beta}{\partial S_0} & \frac{\partial \beta}{\partial E_0} & \frac{\partial \beta}{\partial I_0} & \frac{\partial \beta}{\partial \beta_0} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Observation 2.19. From equations (15), (16) and (17) we get:

$$\begin{cases} \dot{(b_1 + b_2)} = -\gamma b_2 \\ \dot{b_3} = \gamma b_2 - (\lambda + \kappa) b_3 \end{cases}$$

Let us solve the second differential equation. We observe that it is a linear differential equation of first order. Therefore,

$$b_3(t) = c_1 e^{-(\lambda+\kappa)t} + \gamma e^{-(\lambda+\kappa)t} \int_0^t b_2(s) e^{(\lambda+\kappa)s} ds, \text{ where } b_3(0) = 0 \Longrightarrow c_1 = 0$$

Thus,

$$b_3(t) = \gamma e^{-(\lambda+\kappa)t} \int_0^t b_2(s) e^{(\lambda+\kappa)s} ds$$

3. Infection rate β: parameter leading to epidemic waves

Let us recall that the infection rate β is the number of people that an infectious person infects each day and $\beta = pb$ where p is the probability of transmission of the disease and b is the contact rate, the number of people, in average, in contact with an infectious individual.

We have assumed that β is a constant. However, we can expect β to vary in time: decreasing thanks to government restrictions or confinements since these reduce people's contact rate and hence reduce β; and increasing in a lockdown easing period of time. In this section we will see how a SEIR model with a piecewise infection rate $\beta(t)$ representing restrictive measures can lead to epidemic waves. In fact, as we will see in section 4, this can also be achieved by defining the infection rate as a step function and therefore make it easier to adjust the waves to the given data.

3.1 Decreasing and increasing of β depending on the restrictions

Let us contemplate a period of time where the infection rate decreases thanks to restrictions. Let us consider an initial time (in days) $t = 0$ where the infection rate is β_0, a constant. Let $T > 0$ be the day where restrictive measures are taken. These measures are active until day $T + t_m$, where we arrive to a constant infection rate $\beta_1 < \beta_0$. We suppose that this new infection rate β_1 prevails at least until day $T_f > T + t_m$. We could simulate the variation of β in time with the following definition of β:

$$\beta(t) = \begin{cases} \beta_0 & \text{if } 0 \leq t \leq T \\ \beta_0 e^{\frac{-(t-T)}{\zeta}} & \text{if } T < t \leq T + t_m \\ \beta_1 = \beta_0 e^{\frac{-t_m}{\zeta}} & \text{if } T + t_m < t \leq T_f \end{cases} \tag{20}$$

However, in order to model this situation we need to take in account the characteristics of the measures taken. We can distinguish two kinds of measures: hard and brief restrictions; or long and soft restrictions. This is where ζ and t_m take sense. On the one hand, if hard restrictions are taken it makes sense to think that they will last for a brief period before reaching the desired value β_1. On the other hand, if the restrictions are softer, one can think that they will last longer before reaching β_1. This is why in our model we considered two different ζ ($\zeta_1 < \zeta_2$) and t_m ($t_{m_1} < t_{m_2}$) where ζ_1, t_{m_1} correspond to hard and brief restrictions and where ζ_2, t_{m_2} correspond to soft and long restrictions.

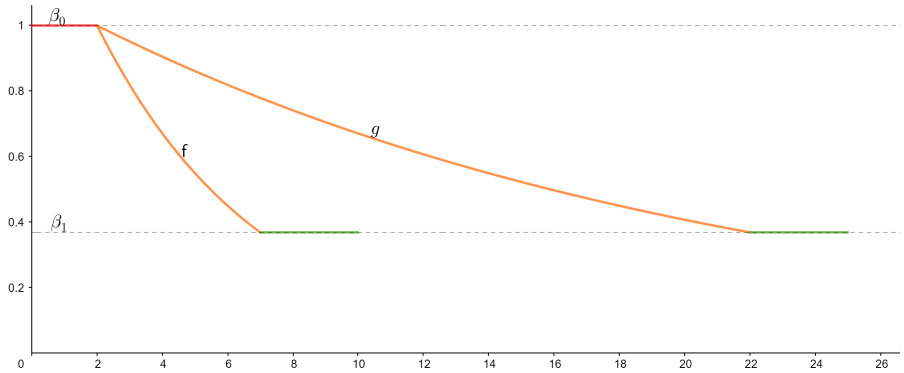


Figure 3: Example of variation of β in time with $\beta_0 = 1$ and $T = 2$. Hard and brief restrictions are represented by function f with parameters $\zeta_1 = 5$, $t_{m_1} = 5$; long and soft restrictions are represented by function g where $\zeta_2 = 20$ and $t_{m_2} = 20$.

Similarly, we can model the increase of the infection rate β. Let us consider an initial time (in days) $t = 0$ where the infection rate is $\beta_1 = constant$. Let $I = [T, T + t_i]$, where $T > 0$ and $t_i > 0$, be a lockdown easing period of time. Let $\beta_2 = constant > \beta_1$ be the final infection rate at time $T + t_i$. We suppose that this new infection rate β_2 prevails at least until day $T_f > T + t_i$. We could simulate the variation of β in time with the following definition of β:

$$\beta(t) = \begin{cases} \beta_1 & \text{if } 0 \leq t \leq T \\ \beta_1 e^{\frac{(t-T)}{\xi}} & \text{if } T < t \leq T + t_i \\ \beta_2 = \beta_1 e^{\frac{t_i}{\xi}} & \text{if } T + t_i < t \leq T_f \end{cases} \tag{21}$$

3.2 SEIR model with infection rate subjected to restrictions

In this section we will numerically show how a piecewise infection rate can lead to epidemics waves.

Following the previous section, we define the infection rate β of our normalised SEIR model (2) subjected to restrictions. We consider first an increase of β followed by a decrease of β and repeat the variations. That is to say, the expression of the first variation of β is:

$$\beta(t) = \begin{cases} \beta_1 & \text{if } 0 \leq t \leq T_1 \\ \beta_1 e^{\frac{(t-T_1)}{\xi}} & \text{if } T_1 < t \leq T + t_i \\ \beta_2 = \beta_1 e^{\frac{t_i}{\xi}} & \text{if } T_1 + t_i < t \leq T_2 \\ \beta_2 e^{\frac{-(t-T_2)}{\zeta}} & \text{if } T_2 < t \leq T_2 + t_m \\ \beta_3 = \beta_2 e^{\frac{-t_m}{\zeta}} & \text{if } T_2 + t_m < t \leq T_f \end{cases} \tag{22}$$

where we imposed $\beta_1 = 0.1$, $T_1 = 0$, $t_i = 30$, $\xi = 50$, $T_2 = 15$ and $t_m = \zeta = \begin{cases} 5 & \text{if hard measures} \\ 30 & \text{if soft measures} \end{cases}$

Remark 3.1. During this study we used Runge-Kutta method with Matlab function *ode*45 in order to solve numerically the classic SEIR model equations (1) and (2). Furthermore, following [2] and [9] and in order to fit best our data in section 4, we imposed the average latent time to be equal to four days, the average infectious period equal to five days and the death rate equal to 3×10^{-3}. Hence, we used parameters $\gamma = \frac{1}{4}$ and $\mu = \lambda + \kappa = \frac{1}{5} + 3 \times 10^{-3}$. These values are frequently used for the COVID19 disease.

After solving numerically equations (2) with the infection rate defined above, we obtained the following plot where we could observe epidemic waves:

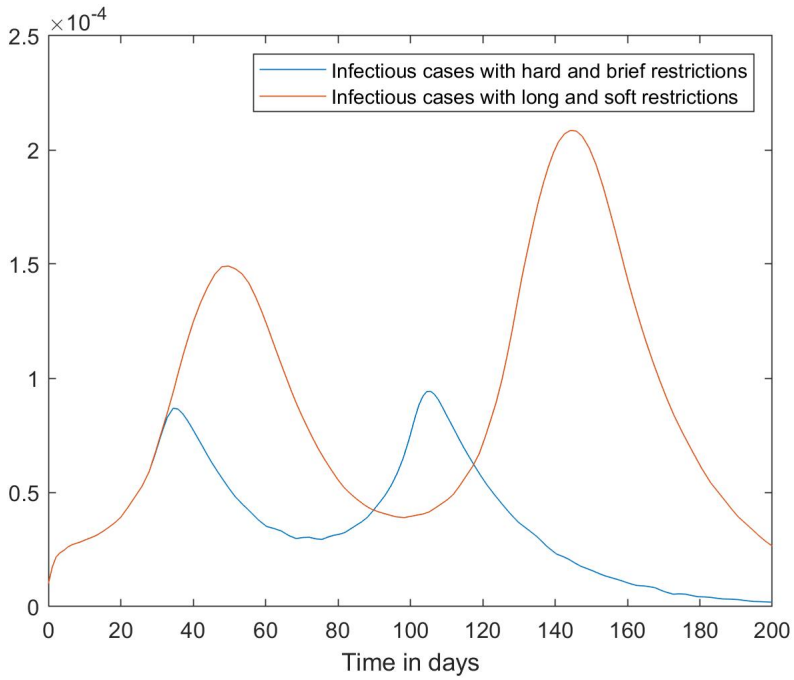


Figure 4: Infectious cases extracted from the normalised SEIR model with an infection rate subjected to restrictions

As we can see, when implementing hard and brief measures the number of infectious cases is considerably lower and lead to shorter waves. This can explain why many countries such as Australia or New Zealand decided to run a zero-COVID policy [10]. This strategy aims to eliminate the virus by keeping its transmission rate as close to zero as possible. In order to accomplish this goal, hard and brief restrictions are taken every time a small number of infectious cases is detected.

However, countries like Spain or France chose to prioritise the health system not to be overwhelmed with long and soft restrictions. Aiming to flatten the infectious curve, these containment strategies lead to higher and longer waves of infectious cases but may end up with an earlier "herd immunity", if this is such an option.

4. Application of the SEIR model to Catalonia's data

Our aim in this section is to reencounter Catalonia's second and third wave of COVID19 infectious people by applying the SEIR model to the data provided by the Generalitat de Catalunya [3].[2] In order to achieve our goal, we will need to find the infection rate β associated to the data. The SEIR model with this infection rate will then, ideally, approximate Catalonia's COVID19 epidemic waves. In contrast with the previous section and due to the fact that the data varies from day to day like a step function, in this section the infection rate will be defined as a step function.

Bearing all this in mind, we propose two different numerical methods to find the infection rate β associated to Catalonia's data. On the one hand, the first numerical method approach is based on the classic and normalised SEIR equations and Hethcote's theorem. On the other hand, the second method is strictly based on the classic and normalised SEIR equations. By means of this second method, we will first analyse the importance of restrictions. Then, we will examine how small variations in the infection rate can lead to higher or smaller epidemic waves. Finally, we will explore the possibility of making predictions.

4.1 First numerical method approach

Our first method is based on Hethcote's theorem (2.16) which says that if $\sigma_0 S(t_0) > 1$, where $\sigma_0 = \beta_0 \frac{1}{\mu}$, and μ is constant, then $E(t) + I(t)$ reaches a (unique) maximum and then decreases. This maximum, reached at $t = T_0$, is equal to

$$E(T_0) + I(T_0) = E(t_0) + I(t_0) + S(t_0) - \frac{1}{\sigma} ln(\sigma S(t_0)) - \frac{1}{\sigma}$$

Let us recall the equations of the classic and normalised SEIR model (2):

$$\begin{cases} \frac{dS(t)}{dt} = -\beta IS & (i) \\ \frac{dE(t)}{dt} = \beta IS - \gamma E & (ii) \\ \frac{dI(t)}{dt} = \gamma E - \mu I & (iii) \\ \frac{dR(t)}{dt} = \mu I & (iv) \end{cases}$$

Adding (ii) and (iii) we observe that

$$(E + I)'(t) = (\beta S(t) - \mu) I(t)$$

At $t = T_0$, $(E + I)'(t) = 0$ and since $I(t) \neq 0$,

$$S(t) = \frac{\mu}{\beta} = \frac{1}{\sigma} \iff \sigma S(t) = 1$$

[2] In order to overcome the weekend effect where less COVID19 tests are done we chose to work with the seven days aggregate data and divided the values by seven. In other words, we defined the daily value of infectious cases as the weekly arithmetic mean.

In our case,

$$S(T_0) = \frac{\mu}{\beta_0} \tag{23}$$

On the other hand, from (i) we get

$$S' = -\beta_0 SI \Longrightarrow \frac{S'}{S} = -\beta_0 I \Longrightarrow \frac{d}{dt}(log(S(t))) = -\beta_0 I(t)$$
$$\Longrightarrow \int_{t_0}^{T_0} \frac{d}{dt}(log(S(t)))dt = -\beta_0 \int_{t_0}^{T_0} I(t)dt$$

We observe that

$$\int_{t_0}^{T_0} \frac{d}{dt}(log(S(t)))dt = log\left(\frac{S(T_0)}{S(t_0)}\right) \overset{(23)}{=} log\left(\frac{\mu}{S_0\beta_0}\right)$$

Thus,

$$log\left(\frac{S_0\beta_0}{\mu}\right) = \beta_0 \int_{t_0}^{T_0} I(t)dt \tag{24}$$

Hence, from (24) we will be able to get an approximation of β_0.

We implemented this method in Matlab by calculating $\int_{t_0}^{T_0} I(t)dt$ first with the Trapezoidal rule, which states that:

$$\int_a^b f(t)dt \approx (b-a)\frac{f(a)+f(b)}{2} \tag{25}$$

Since we wanted our method to be more efficient, we then changed to Simpson's rule:

$$\int_a^b f(t)dt \approx \frac{(b-a)}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right] \tag{26}$$

Remark 4.1. Simpson's rule is exact for polynomials of degree $n \leq 2$ whereas Trapezoidal rule is exact for polynomials of degree $n \leq 1$. Moreover, the global error of the composite Simpson's rule and the composite Trapezoidal rule are, respectively, of order h^4 and h^2 where $h = (b-a)/2$. Note that we used the composite rules in order to be more accurate. Hence, Simpson's rule is more efficient since it achieves a given level of accuracy faster and is exact for second degree polynomials.

Once we had calculated the integral of infectious people, we solved equation (24) with the Newton-Raphson method. We obtained then an approximation of the infection rate β_0 at the beginning of the wave. We put that value of β in our SEIR equations and we used Matlab's function *ode*45 to solve the system.

In order to verify the proper functioning of this method, we first tried it on a data created by a known SEIR. That is to say, we calculated the number of infectious people per day with our SEIR equations where we imposed a value for β and then tried to reencounter that value of β with our numerical method. The result was satisfactory since we reencountered the initial β with a small error (magnitude of 10^{-5}).

We then applied our method to Catalonia's data. The resulting number of infectious people per day was however considerably different from the data. This was due to the fact that we were not considering that the peak observed in the data was altered by the government's restrictions. In other words, we were supposing that the observed peak corresponded to the theoretical peak of a SEIR model whose value is known (5), whereas in reality the observed peak is only a fraction of the wave's ascent. If the government had not imposed restrictive measures the "real" or theoretical peak would have been reached later in time.

Let us illustrate this with the following figure:

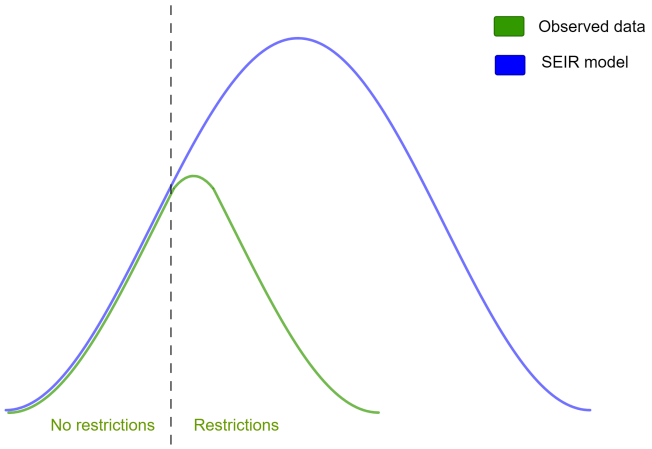


Figure 5: Schema illustrating the differences between waves of infectious cases with the observed data, which is subject to restrictions, and the theoretical SEIR model

4.2 Second numerical method

Since our first method was not accurate to model the real data, we came up with a second method. This new method is strictly based on the classic and normalised SEIR model equations (2):

$$\begin{cases} \frac{dS(t)}{dt} = -\beta IS & (i) \\ \frac{dE(t)}{dt} = \beta IS - \gamma E & (ii) \\ \frac{dI(t)}{dt} = \gamma E - \mu I & (iii) \\ \frac{dR(t)}{dt} = \mu I & (iv) \end{cases}$$

From (iii) we obtain an equation for β:

$$\begin{aligned} I''(t) = \gamma E' - \mu I' &\overset{(ii)}{=} \gamma\beta IS - \gamma^2 E - \mu I' \overset{(iii)}{=} \gamma\beta IS - \gamma(I' + \mu I) - \mu I' \\ \implies \beta &= \frac{I'' + (\gamma + \mu)I' + \gamma\mu I}{\gamma IS} \end{aligned} \tag{27}$$

However, this equation depends on S, the susceptible number of people, which is difficult to approximate at $t \neq t_0$. Therefore, we studied equation (i) and obtained equality (28):

$$\begin{aligned} \frac{S'}{S} = -\beta I \implies \int_{t_0}^{T} \frac{S'}{S} dt = -\beta \int_{t_0}^{T} I dt &\implies log(S(T)) = log(S(t_0)) - \beta \int_{t_0}^{T} I dt \\ &\implies S(T) = S(t_0) e^{-\beta \int_{t_0}^{T} I dt} \end{aligned} \tag{28}$$

Hence, for $T \geq t_0$,

$$\beta(T) = \frac{I''(T) + (\gamma + \mu)I'(T) + \gamma\mu I(T)}{\gamma IS(t_0) e^{-\beta \int_{t_0}^{T} I dt}} \tag{29}$$

We then proceed to approximate β numerically with this new equation (29). As seen in the previous section, we start by calculating $\int_{t_0}^{T_0} I(t)dt$ first with the Trapezoidal rule (25) and then with Simpson's rule (26). Afterwards, we approximate the first and second derivatives of I with the symmetric difference quotient. Recall that the symmetric difference quotient for the first derivative is:

$$f'(t) \approx \frac{f(t+h) - f(t-h)}{2h} \tag{30}$$

and for the second derivative is:

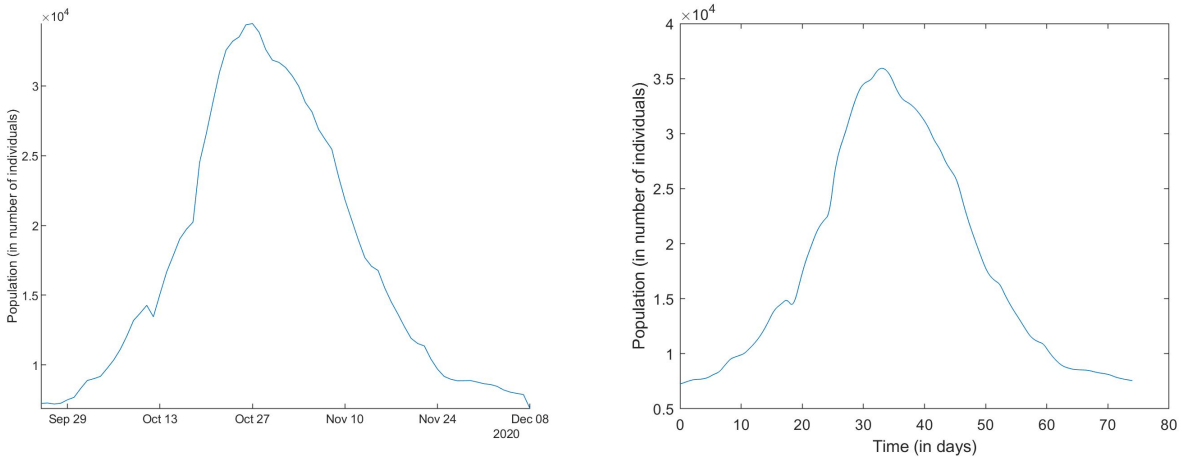
$$f''(t) \approx \frac{f(t+h) - 2f(t) + f(t-h)}{h^2} \tag{31}$$

Finally, we solve equation (29) with the Newton-Raphson method.

As before, we first verified the proper functioning of this new method by trying it on a data created by a known SEIR. The test was successful with almost all values of β, but some of them returned wrong values.

We solved the problem by changing the Newton-Raphson method to the Regula falsi method. Indeed, when running the Newton-Raphson method, the derivative of one of the iterations was close to zero and caused the method to fail. With this modification all the tests were successful.

We then applied our method to Catalonia's data and calculated $\beta(t)$ every day of the second wave (25/05/2020 to 08/12/2020). When we introduced the found vector $\beta = (\beta(1), ..., \beta(t_f))$ in our code to solve the SEIR equations, we obtained really similar waves:

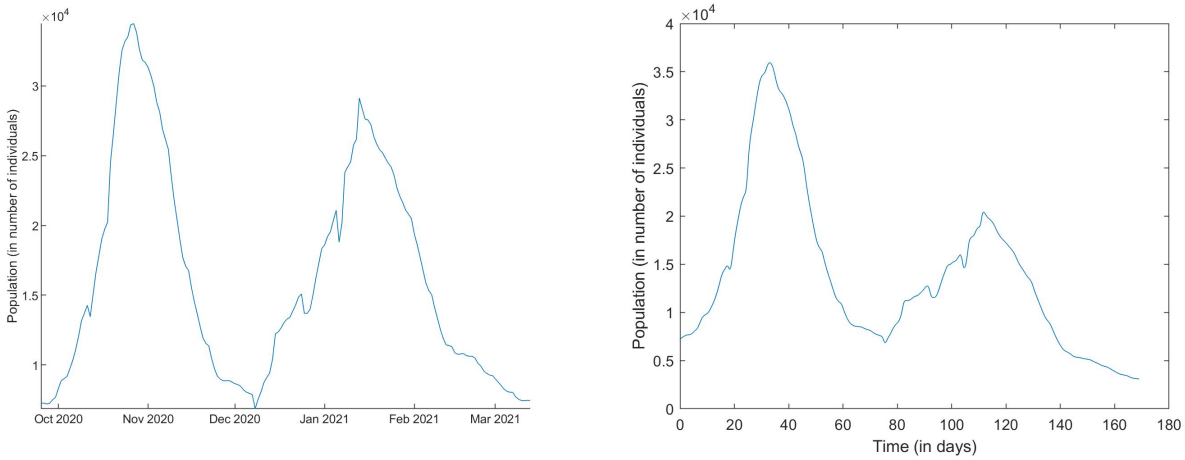


(a) Plot extracted from the data

(b) Plot obtained by the SEIR model

Figure 6: Daily confirmed cases during the second wave according to the data or the SEIR model where the day number 0 in figure 6b corresponds to September 25th, i.e the first day of figure 6a

We repeated this process adding the third wave (08/12/2020 to 13/03/2021) and observed again similar waves:



(a) Plot extracted from the data

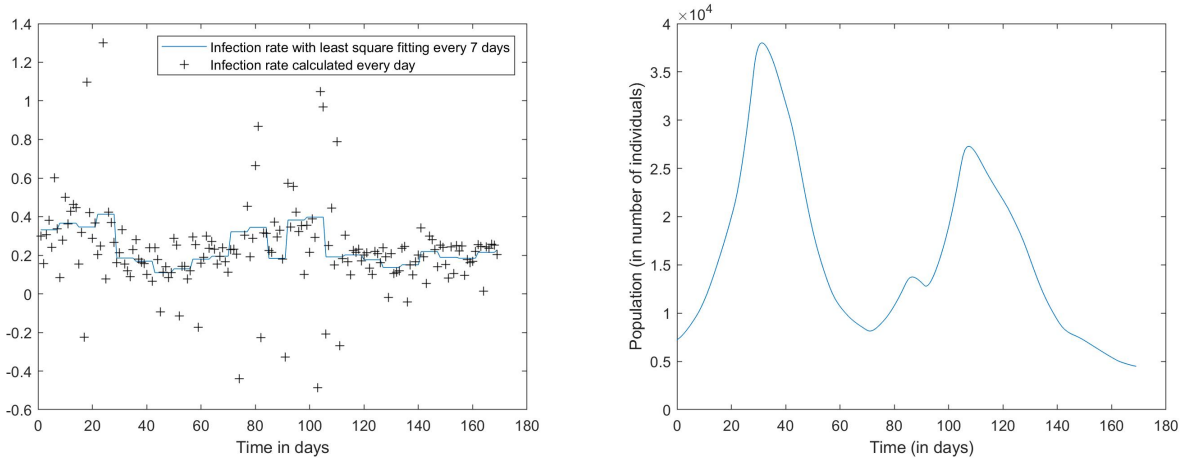
(b) Plot obtained by the SEIR model

Figure 7: Daily confirmed cases during the second and third wave according to the data or the SEIR model

As we can see, the height of the simulation's second peak is less accurate. However, this difference will be reduced in the next steps.

Certain values $\beta(t)$ seemed to vary a lot from one day to another. In addition, we found some negative values of β. This was due to the fact that we were "cheating" in some way by choosing β so that when integrating we would find exactly the given data. That is the reason why we decided to impose our β to be a constant value within a given interval of time by performing a least square fitting.

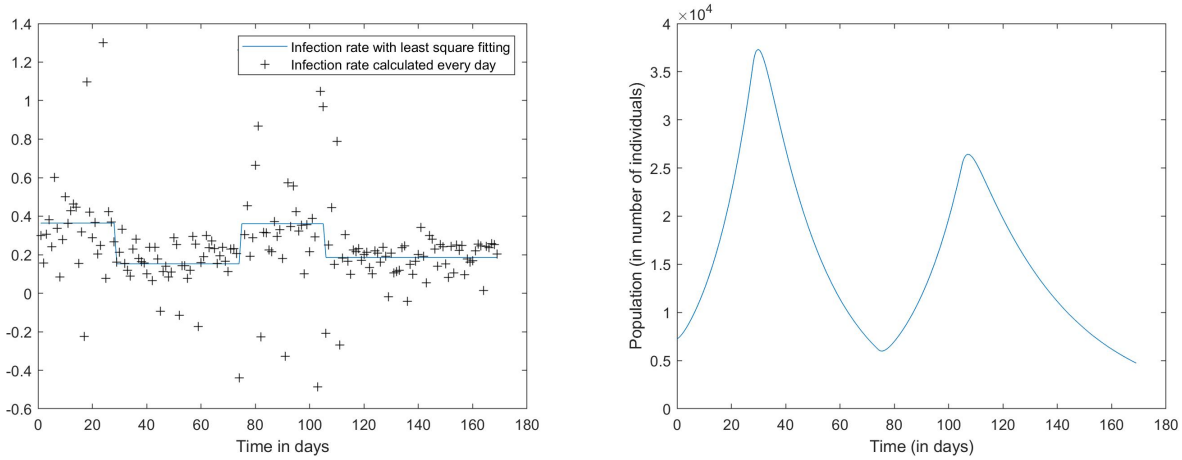
We started by considering intervals of seven days in order to have a weekly value of the infection rate. We obtained the following plots for the second and third wave:



(a) Infection rates β before and after least square fitting (b) Daily infectious cases extracted from the SEIR model

Figure 8: Modelling of the second and third wave with least square fitting of the infection rate every seven days

While observing these graphics and their accuracy to model both waves, we asked ourselves if we could obtain a similar plot with just four different infection rates $(\beta_1, \beta_2, \beta_3, \beta_4)$, one for each variation of the infectious curve.

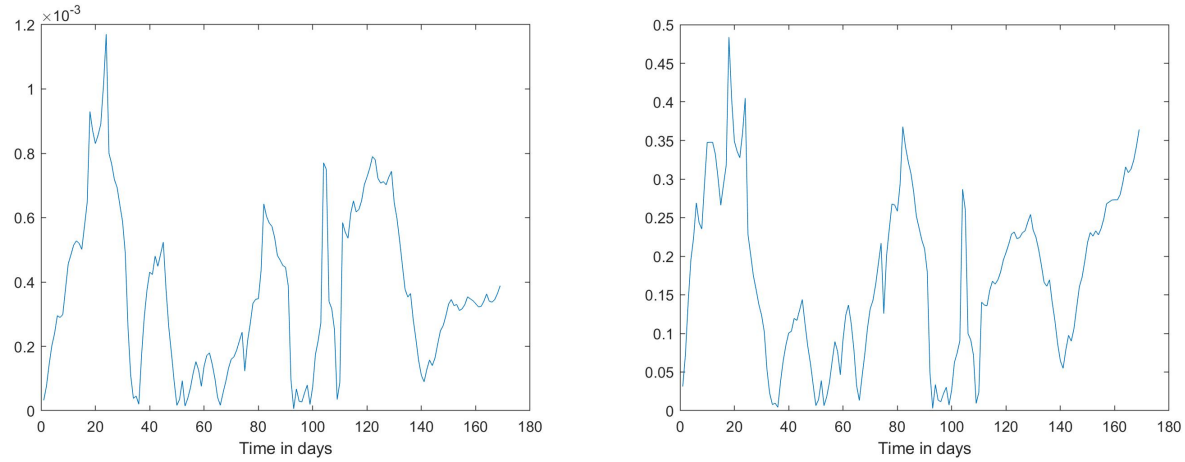


(a) Infection rates β before and after least square fitting (b) Daily infectious cases extracted from the SEIR model

Figure 9: Modelling of the second and third wave with four infection rates obtained with least square fitting

We need to remark that this new solution of the infectious cases in time presents a higher error approximation when comparing with the real data as we can see in figure 10. Moreover, the peaks detected in the SEIR model have a delay of two or three days from the "real" peaks.

Nonetheless, the solution obtained is smoother than the previous ones and in fact is $\mathcal{C}^\infty$. Furthermore, we cannot affirm that the peaks observed in the real data correspond to the real situation since we need to take into account some delay caused by weekends, bank holidays, required days to determine the result of COVID19 tests, etc. Thus we can say that this indeterminacy is reasonable.



(a) Absolute error

(b) Relative error

Figure 10: Error approximation when modelling the second and third wave with four βs obtained with least square fitting

Even though we decided almost arbitrarily the dates where we wanted to change the value of β, we observed that they could actually be associated to the dates where restrictive measures were taken in Catalonia. Indeed, according to AQUAS's report [4] new restrictions were imposed on November 15th, 25th and 30th in order to minimise the impact of the second wave and January 7th for the third wave. In our model, we changed the value of β on November 23rd, December 8th and January 8th. Note that December 8th corresponded to the long weekend of the Constitution, were mobility was allowed in Catalonia. Hence, we can say that restrictions clearly determine the value of β.

4.3 Analysis of the importance of restrictions

In the previous section we observed a relationship between the value of the infection rate β and the days where restrictive measures were taken. In this section we will continue to study the importance of restrictions with regards to a wave's impact in a population. We will consider as a basis the SEIR model with four different infection rates (figure 9b).

Firstly, we asked ourselves what could have happened if no restrictive measures had been taken during the second wave. In other words, what shape would the infectious curve have if the initial value of the infection rate β had not changed over time. We modelled this situation and were able to illustrate it in the following figure, where we can see the differences between the infectious curves of the basis model with four infection rates (in blue) and the possible scenario without restrictions (in red) during the period of time corresponding to the second and third wave.

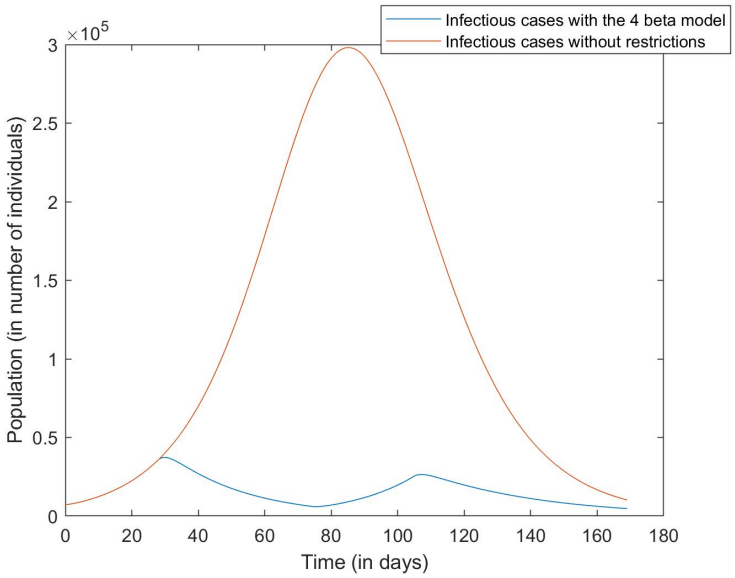


Figure 11: Daily infectious cases extracted from the SEIR model with or without restrictions during the period of time corresponding to the second and third wave

As we can see in figure 11, the difference of infectious cases is impressive as the peaks differ by more than $2 \cdot 10^5$ cases, over a population around $7.5 \cdot 10^6$. This difference could have prompted to a vast number of hospitalizations and deaths, collapsing the health system. Hence, restrictive measures are essential for a population not to enter a critical situation.

Remark 4.2. When modelling the SEIR without restrictions, we observed that the peak of the exposed and infectious' sum corresponded to the value of Hethcote's theorem (5) as we can see in figure 12 below.

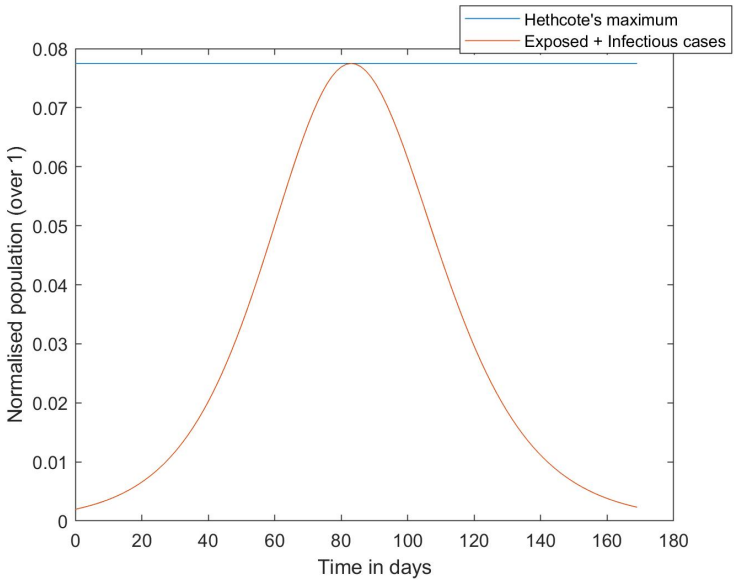
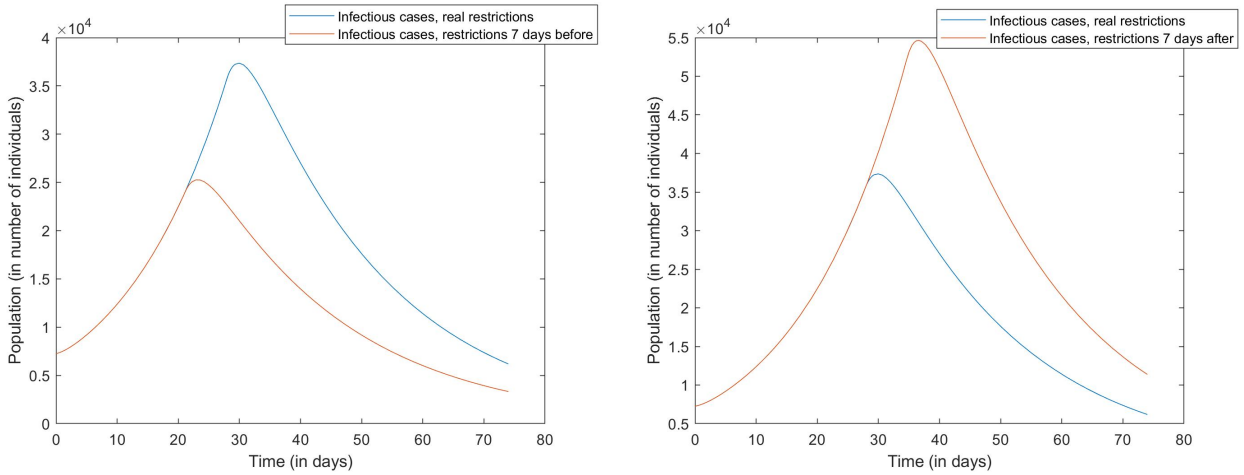


Figure 12: Hethcote's maximum of the normalized SEIR without restrictions

Once we had calculated the four values of the infection rate overtime we asked ourselves what would have happened if the restrictive measures had been taken a few days before or a few days after.

With that in mind, we started by moving the date where the first infection rate β_1 changed to β_2.



(a) Restrictions taken seven days before

(b) Restrictions taken seven days after

Figure 13: Daily infectious cases during the second wave extracted from the SEIR model when varying the date where restrictions are taken

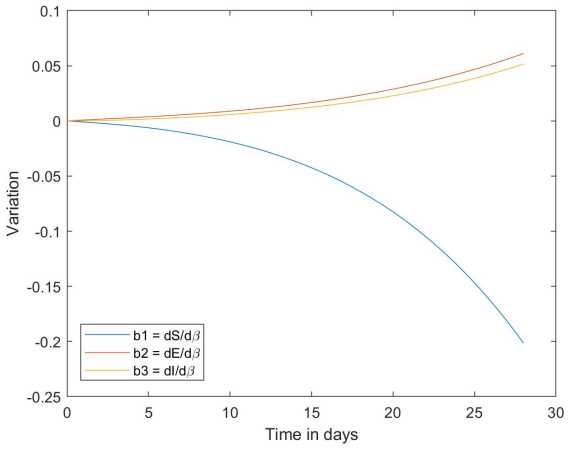
On the one hand, if we had introduced restrictive measures a week before we would have had only $2/3$ of the infectious cases. In addition, it would have taken approximately twenty days less for the wave to return to its initial value. On the other hand, if these measures had been taken a week after, the infectious cases would have increased by a factor of $4/3$ and the wave would have lasted for a longer period of time.

We can conclude that it is of high importance to introduce restrictive measures rapidly whenever we detect several new cases in a population, at least from a medical point of view.

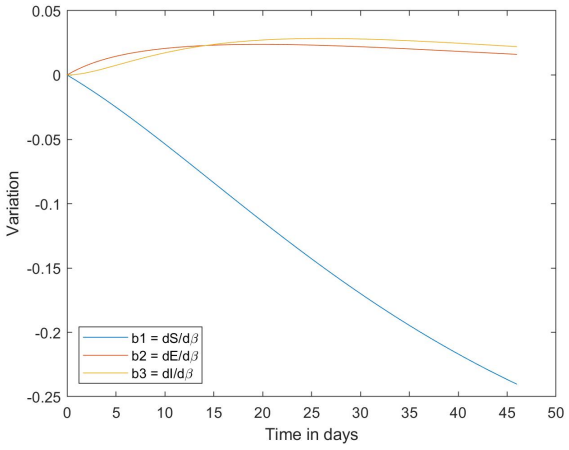
4.4 Implementing variationals

All this time, we have considered as the basis model the SEIR model with four different infection rates (figure 9b). However, the calculations of the values of the different infection rates can experience a certain numerical approximation error which can influence the number of Susceptible, Exposed or Infectious people over time. In addition, this influence phenomena could also be caused by small variations in the restrictive measures. These are some of the reasons why we decided to study the variationals regarding parameter β.

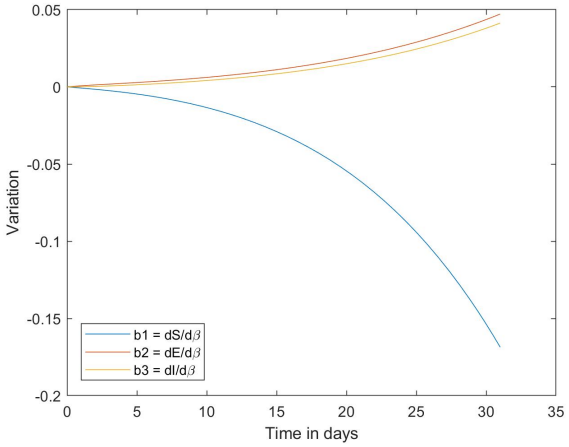
After solving numerically the variationals equations regarding parameter β (19), we obtained the following plots:



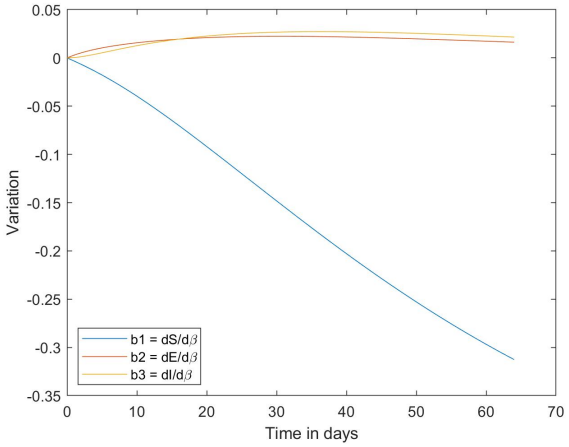
(a) Variation regarding β_1: the rise of the second wave



(b) Variation regarding β_2: the fall of the second wave



(c) Variation regarding β_3: the rise of the third wave



(d) Variation regarding β_4: the fall of the third wave

Figure 14: Variationals regarding parameter β

Thanks to these plots we can determine how an initial numerical approximation error of the infection rate will influence our SEIR model. Let us focus on the infectious group $I(t)$. Let ϵ_1 be the difference between our approximation of the first infection rate β_1 and the real value. Let t_{max} be the time where b_3 is maximal and t_{end} be the final time. We observe that at $t_{max} = t_{end}$, b_3 is maximal and approximately equal to 0.0516 (figure 14a). Hence, at the peak of the second wave the predicted infectious cases will

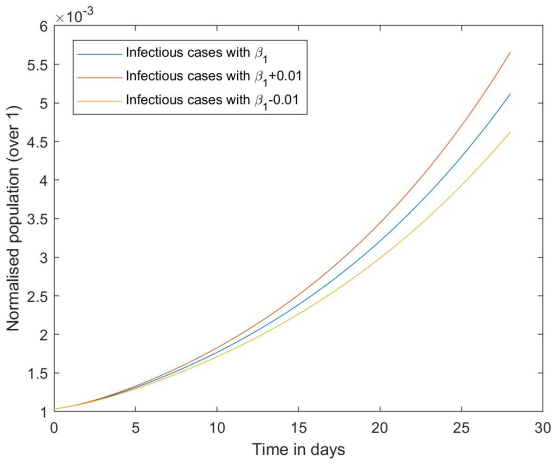
differ by $\bar{\epsilon}_1 \approx 0.0516 \times \epsilon_1$ from the "real" ones. Similarly, we can find approximations for the maximal and final differences of infectious cases associated to every infection rate. We did so and gathered the results in the table below.

Infection rate	Initial difference	Final difference at t_{end}	Maximal difference held at t_{max}	t_{max}
β_1	ϵ_1	$0.0516 \times \epsilon_1$	$0.0516 \times \epsilon_1$	28
β_2	ϵ_2	$0.0221 \times \epsilon_2$	$0.0284 \times \epsilon_2$	26
β_3	ϵ_3	$0.0413 \times \epsilon_3$	$0.0413 \times \epsilon_3$	31
β_4	ϵ_4	$0.0216 \times \epsilon_4$	$0.0272 \times \epsilon_4$	36

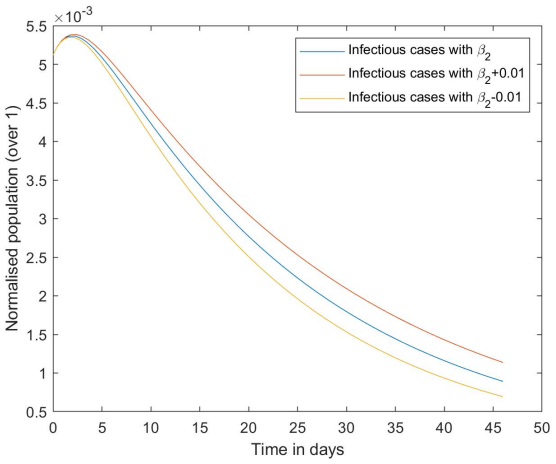
Figure 15: Initial, final and maximal difference of every infection rate

We could also interpret the initial difference ϵ as the result of small variations in the restrictive measures. Therefore, at time t_{max} is where we would have detected a greater variation in the infectious number of cases if we had taken different restrictive measures, either a little bit harder or softer. In order to numerically confirm this we changed the value of β_1 and β_2 by adding or subtracting 0.01.

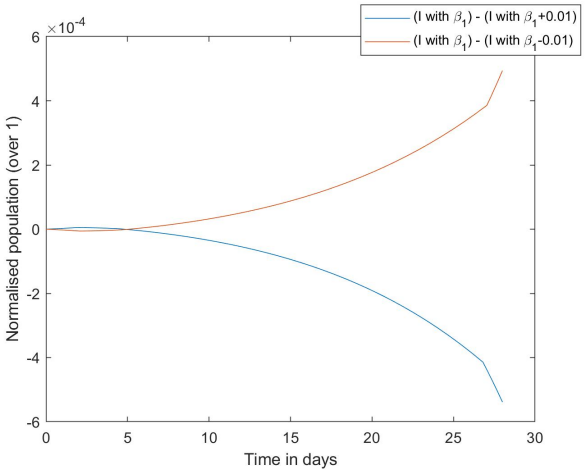
Indeed, as we can see in figures 16c and 16d below, the maximal difference is held at the respective times t_{max} found at table 15. Furthermore, we see that at t_{max} the infectious cases differ by approximately 5.1×10^{-4} cases for β_1 and 2.2×10^{-4} cases for β_2. Since our initial difference was $\epsilon_1 = 0.01$, we expected our maximal differences to be respectively $\bar{\epsilon_1} \approx 0.0516 \times 0.01 = 5.16 \times 10^{-4}$ and $\bar{\epsilon_2} \approx 0.0221 \times 0.01 = 2.21 \times 10^{-4}$. Hence, we obtain the results from table 15.



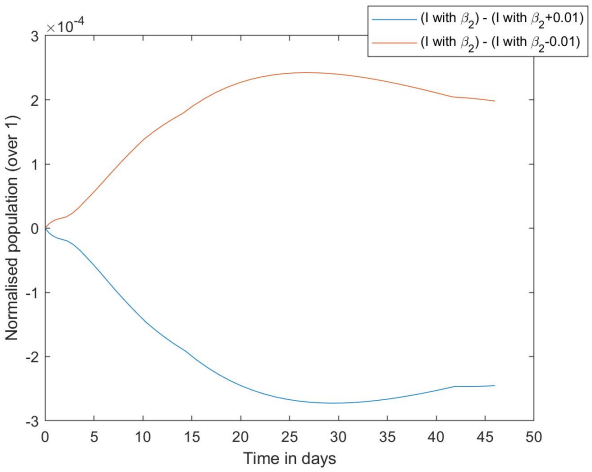
(a) Daily infectious cases when varying β_1 by 0.01



(b) Daily infectious cases when varying β_2 by 0.01



(c) Difference between curves from figure 16a



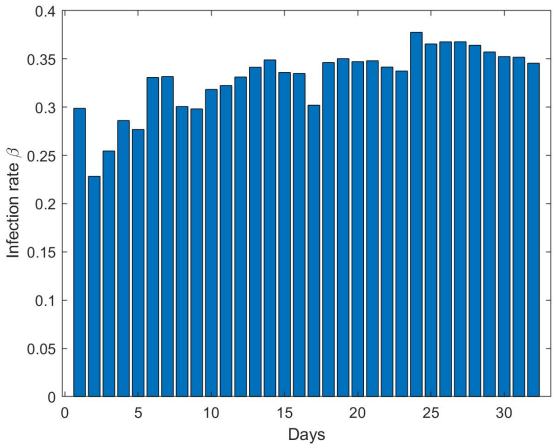
(d) Difference between curves from figure 16b

Figure 16: Variation of daily infectious cases when varying β by 0.01

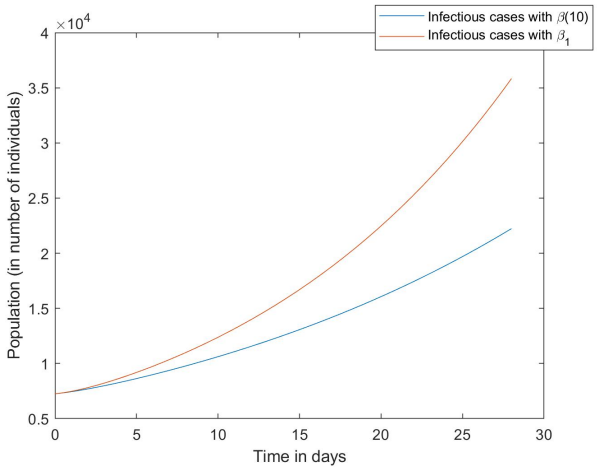
4.5 Can we do predictions?

When calculating the four values of the infection rate, we asked ourselves whether it was necessary to know the value of the infection rate β for each day before doing a least square fitting or if with just the first ten values of β (for example) we could obtain a similar SEIR model. If the second option were to be accurate we could then do predictions.

With that in mind, we calculated the accumulated values of the infection rate for each day. That is to say, the value of β with least square fitting in the interval of time $[1, T]$ for each day T of the second wave's rise. The aim was to see if there was a day where the values became steady. However, as we can see in figure 17a, there seemed not to be a day where the values stabilised. This observation became more clear when modelling the rise of the second wave with the accumulated infection rate at day 10 and the infection rate β_1 (figure 17b).



(a) Daily accumulated value of the infection rate



(b) Daily infectious cases during the rise of the second wave extracted from the SEIR model with $\beta(10)$ and β_1

Figure 17: Can we do predictions?

Indeed, at the end of the rise the infectious cases differ from more than 10^4 cases.

5. Conclusions

The goal of this Bachelor's degree thesis was to study the dynamics of epidemic waves in a SEIR model, focusing on the COVID19 pandemic and taking Catalonia as an example. Our main finding is that restrictive measures and lockdown easing periods influence COVID19's wave-like behaviour. Hard restrictions are more effective than softer ones; and it is also vital to impose them as soon as new cases are detected to avoid the collapse of our health system and a large number of deaths.

In terms of approach, this thesis has found a pretty accurate numerical method to apply a simple SEIR model to real data. In addition, thanks to the model's simplicity, we were able to apply diverse theoretical results previously examined like Hethcote's theorem or the variationals equations. These results enabled us to analyse the importance of restrictions and to study how small variations in the infection rate could influence the number of infectious cases over time.

With regards to the infection rate, we saw how its definition was key to the development of epidemic waves. Indeed, changing its value by increasing and decreasing it over time causes the SEIR model to exhibit several infectious waves instead of only one. As a matter of fact, we only need two different infection rates $\beta_1 > \beta_2$ to model an infectious wave shaped by restrictive measures.

All in all, the aim of this thesis has been to present an original numerical method to analyse the dynamics of epidemic waves of infectious diseases such as COVID19 and draw conclusions on how best to tackle them. Given its scope, this thesis has compared "hard" and "soft" restrictions, without distinguishing each of these individually - for example, wearing a face-mask compared to closing restaurants. The list of "hard" and "soft" measures is indeed very long and countries around the world have chosen to impose different ones at different times. For example, countries like Belgium imposed a "social bubble" rule, closed restaurants for months but left mobility free, while in Spain mobility between regions was forbidden but restaurants remained open (with rules varying significantly between regions). It would be interesting to use this model to compare the impact of these different restrictions on COVID19 wave-like behaviour and assess their respective usefulness. Furthermore, it would also be interesting to explore the possibility of making predictions, maybe without using all the data. This thesis encourages further research in these directions.

6. Bibliography

References

[1] Agencias. Nueva formulación aporta datos más fiables sobre la evaluación de la pandemia. https://www.lavanguardia.com/vida/20210204/6222346/nueva-formulacion-aporta-datos-mas-fiables-sobre-evaluacion-pandemia.html. *La Vanguardia*, 2021.

[2] Nidhal ben Khedher, Lioua Kolsi, and Haitham Alsaif. A multi-stage SEIR model to predict the potential of a new covid-19 wave in KSA after lifting all travel restrictions. *Alexandria Engineering Journal*, 60(4):3965–3974, 2021.

[3] Generalitat de Catalunya. Salut/Dades COVID. Població total, agregació 7 dies. https://dadescovid.cat/descarregues, 2021.

[4] Generalitat de Catalunya, AQUAS, CMCiB, and BIOCOMS. Evolució dels casos i de la Rt del SARS-CoV2. https://dlscitizens.blob.core.windows.net/rtreports/archived/20210611/CAT/InformeCasosRt_CAT_09.pdf, 2021.

[5] Odo Diekmann, Johan Andre Peter Heesterbeek, and Johan AJ Metz. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *Journal of mathematical biology*, 28(4):365–382, 1990.

[6] Alberto Godio, Francesca Pace, and Andrea Vergnano. SEIR modeling of the italian epidemic of SARS-CoV-2 using computational swarm intelligence. *International Journal of Environmental Research and Public Health*, 17(10):3535, 2020.

[7] Herbert W Hethcote. The basic epidemiology models: models, expressions for R0, parameter estimation, and applications. In *Mathematical understanding of infectious disease dynamics*, pages 1–61. World Scientific, 2009.

[8] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

[9] Stephen M Kissler, Christine Tedijanto, Edward Goldstein, Yonatan H Grad, and Marc Lipsitch. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science*, 368(6493):860–868, 2020.

[10] Anna Llupià, Israel Rodríguez-Giralt, Anna Fité, Lola Álamo, Laura de la Torre, Ana Redondo, Mar Callau, and Caterina Guinovart. What is a Zero-COVID strategy and how can it help us minimise the impact of the pandemic? *ISGlobal*, 26(1), 2020.

[11] David M Morens and Jefferey K Taubenberger. 1918 Influenza: The mother of all pandemics. *Emerging Infectious Diseases*, 12(1):15–22, 2006.

[12] World Health Organization. WHO COVID-19 Dashboard, 2020. Available online: https://covid19.who.int/. Last accessed on June 2021.

[13] G Sallet. Inria & ird epicasa09 avril 2010. *Rapport technique*, 2010.

[14] Pauline Van den Driessche and James Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical biosciences*, 180(1-2):29–48, 2002.