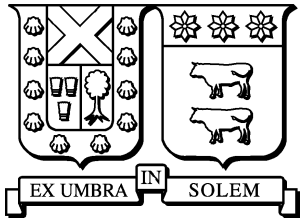
FEDERICO SANTA MARIA TECHNICAL UNIVERSITY
INFORMATICS DEPARTMENT
SANTIAGO, CHILE



A HEURISTIC APPROACH FOR DETERMINING AN EFFICIENT VACCINATION PLAN UNDER A SARS-CoV-2 EPIDEMIC MODEL

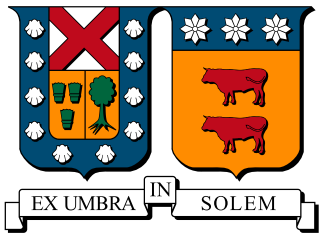
Thesis to apply for the title of
BACHELOR OF COMPUTER SCIENCE

Author
Claudia Hazard Valdés

Thesis adviser
Elizabeth Montero Ureta
Thesis co-advisor
José Luis Martí Lara

March 29, 2021

FEDERICO SANTA MARIA TECHNICAL UNIVERSITY
INFORMATICS DEPARTMENT
SANTIAGO, CHILE



A HEURISTIC APPROACH FOR DETERMINING AN EFFICIENT VACCINATION PLAN UNDER A SARS-CoV-2 EPIDEMIC MODEL

Thesis to apply for the title of
BACHELOR OF COMPUTER SCIENCE

Author
Claudia Hazard Valdés

Thesis adviser
Elizabeth Montero Ureta

Thesis co-adviser
José Luis Martí Lara

March 29, 2021

Acknowledgements

I would like to express my gratitude to my thesis adviser Elizabeth Montero for her guidance, advises and the time she spent supporting my work. Also a special gratitude to my parents and brothers for the support and for helping me to keep my mind clear and to do not stress so much. To all the people I meet in my university time from the first to this last year. Those that help me to keep striving. To the people I meet during my internship who helped me with languages and to have a more open mind. And finally, to all my friends who support me every time I was down or stressed. Without these people nothing would have been the same.

Abstract

The COVID-19 pandemic produced by SARS-CoV-2 virus has generated big changes all over the world. Although, to date vaccinations are being carried out with a limited quantity per country making necessary the design of proper distribution plans. For other diseases such as Influenza or Ebola, research has been carried out in the best way to vaccinate the population. With this background and the fact that COVID-19 slowed down the worldwide economy and mobility, an algorithm is being sought to search for the best vaccine allocation, decreasing the number of infected people.

Keywords: *Vaccination plan, SARS-CoV-2, Optimization, Heuristics, Epidemiological Model.*

Contents

Introduction

Several different viruses have been found in the history of humanity. Some of them, which trespass from animals to humans, generate epidemics like Influenza, HIV, Ebola and currently COVID-19.

The COVID-19 disease is caused by the SARS-CoV-2 virus, which corresponds to the seventh member of the range of corona-viruses that have been infected humans. Its first appearance was recorded in December 2019 in Wuhan, China [19]. The symptoms of the first patients were fever, general discomfort, dry cough and dyspnea. For this reason it was first diagnosed as viral pneumonia. After four months of the first recorded case in China, the virus spread rapidly to many other countries, thus, on March 11 *World Health Organization (WHO)* it was declared as a pandemic, the fifth since 1918 (H1N1).

Currently, worldwide 120,745,792 infected have been detected and 2,671,764 deaths, thereof in Chile 905,212 infected have been confirmed. The large number of infected in a short period of time has caused the collapse of most health systems in the world. Since the number of infected becomes bigger than the number of patients, the medical staff and infrastructure can withstand, having to decide which patients to treat first, generating not only deaths, but also psychological effect on the population and even more on the medical staff who must deal directly with this.

Due to the current global connectivity and the virus contagiousness, it is assumed that one infected person infects two or three people. Each country has set

different measures to decrease viral spread and increase social distance. Among the measures taken, borders have been closed, quarantines have been carried out or activities have been limited to a certain number of people. All those measures generated changes like working remotely, not only in companies but also at educational level. Regarding transportation, a reduction of flights of different airlines along with decreased movement in public transport occurred. Economically, several companies got insolvent due to a lack of clients, bringing along a higher rate of unemployment.

Since COVID-19 is classified as a pandemic, scientists from different countries have been working on vaccines for it. Currently, in Chile, there are commercial agreements to obtain four types of vaccines this are, Sinovac, Pfizer–Biontech, Janssen and Astra Zeneca. Despite of how fast vaccines production is going it will not be sufficient to supply quickly all countries. This, in addition to the high cost associated to their production. For this reason, the need arises to determine the best vaccination plan for SARS-CoV-2 considering a limited number of vaccines. Thus, reducing the number of people infected over time for a virus which there is limited information and a new vaccine whose effectiveness will not be fully credited.

From all the context mentioned above, a method for determining an efficient vaccination plan for SARS-CoV-2 will be presented in this study. This approach, will be based on previous vaccination models. Also, adaptations of previews epidemic models will be studied, in a way to model SARS-CoV-2 spread the closest to reality.

State of the Art

Through the years, most disease spreads have been modeled in a variety of ways, taking into account various key factors. In this chapter, it is shown how some respiratory disease spread can be modeled, which factors have been used and how models with and without vaccination have been implemented.

First, we explain what is an epidemic model and mention some of the most known models with their parameters. Secondly, it is exposed three different approaches using mathematical programming or computer simulation. Finally, the state of the art for spread disease models with vaccination is deepened.

4.1 Epidemic Models

Epidemics models are tools able to describe how a disease spreads through the population. Thus, they can be used to predict the outbreak and the effect of different methods for controlling, decreasing or delaying disease spread.

These epidemic models are generally solved as ordinary differential equations (ODEs), but also, in some complicated cases can be solved by partial differential equations (PDEs) [25].

Epidemic models can be classified according to the steps of their spreading process. In the following we explain one of the most well known spreading disease model, the SIR Model. Moreover, we also detail some of their main variations.

4.1.1 SIR Epidemic Model

The *Susceptible-Infected-Recovered Epidemic Model* was proposed by McKendrick and Kermack in 1927 [1] and is one of the simplest approaches for modeling disease spreading. This model divides the population into three groups:

- Susceptible (S): Healthy individuals who can catch the disease. Number of susceptible individuals.
- Infected (I): Individuals who currently have the disease and can transmit it to susceptible individuals. Number of infected individuals.
- Recovered (R): Individuals who had the disease and now are immune. Number of recovered individuals.

For the SIR model the following assumptions are made. First, the sum of individuals in these three groups corresponds to the population size (N). In this way the model is closed, entailing that the population size is always the same.Secondly, the contact rate (contact between individuals) per unit of time will be β. Third, after contact with an infected individual, a susceptible individual gets infected with rate C_e and becomes immediately infected (there is not incubation period). Fourth, infected individuals recover after r delay of time or recovery rate as $1/r$ (length which a person remain infected).

Based on all these assumptions, the SIR system can be described by the following equations:

Equation 4.1 establishes that the population is closed.

$$N = S(t) + I(t) + R(t) \tag{4.1}$$

Equation 4.2 shows that the number of susceptible individuals decreases in time according to the contact rate.

$$\frac{dS}{dt} = -\beta SI \tag{4.2}$$

Equation 4.3 indicates that infected individuals are increasing according to the contact rate and decreasing after recovery.

$$\frac{dI}{dt} = \beta SI - \frac{I}{r} \tag{4.3}$$

Equation 4.4 shows that the number of recovered individuals increases according of the number of infected individuals and the recovery delay.

$$\frac{dR}{dt} = \frac{I}{r} \tag{4.4}$$

An indicator of the transmission capacity of a virus is settled by the basic reproduction number (R_0). It can be obtained as $R_0 = \beta N r$ or $R_0 = C_e r$ and represents the average number of infections generated by a single person through the population. Different diseases, that share the same infectious phases can be modeled by knowing the values of the parameters mentioned above.

The SIR model and its equations can be represented by the flow diagram shown in figure 4.1. The population is divided into susceptible (S), infected (I), and recovered (R). The susceptible population gets infected with the rate β and the quantity of infected population that gets infected corresponds to βSI. Infected individuals get recovered with a rate $1/r$ with r as the number of days to get recovered implying that I/r infected people get recovered in time.

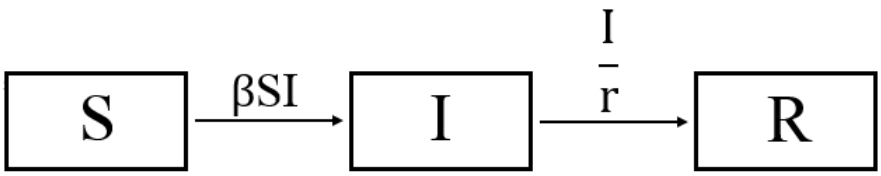


Figure 4.1: SIR model flow diagram.

Adding b as birth rate, d as death rate and a as the disease death rate, are shown in the next equations:

$$\frac{dS}{dt} = bN - dS - \beta SI \tag{4.5}$$

$$\frac{dI}{dt} = \beta SI - \frac{I}{r} - (d + a)I \tag{4.6}$$

$$\frac{dR}{dt} = \frac{I}{r} - dR \tag{4.7}$$

4.1.2 SEIR epidemic model

The *Susceptible-Exposed-Infected-Recovered Epidemic Model* is a variation of the SIR Model [18]. A new group of *exposed* individuals is incorporated in this case. Referring to some diseases, after contact between a susceptible and an infected person, the susceptible individual becomes exposed. In this state, the person is infected but not infectious (can not infect susceptible individuals) because of the incubation period of the disease.

Once included this new group, the new factor e represents the rate of exposed individuals getting infectious and the equation of infected should be updated as shown in equation 4.8.

$$\frac{dE}{dt} = \beta SI - Ee \tag{4.8}$$

Moreover, individuals get exposed with the rate e as shown in equation 4.9.

$$\frac{dI}{dt} = Ee - \frac{I}{r} \tag{4.9}$$

The number of infected individuals depends on the exposed individuals and the rate of getting infectious.

To bring the models closer to reality, birth and death rates are incorporated into these models. In our study, given that the study of the virus considers limited time ranges, it is assumed that these factors are the same to maintain the same population size over time.

4.1.3 SIQR epidemic model

Another variation of the *SIR Epidemic Model* is the *Susceptible-Infected-Quarantined-Recovered Epidemic Model (SIQR)*. In this model, some of the infected individuals go on quarantine with a rate q and stop infecting susceptible individuals.

The number of infected individuals in this case is given by equation 4.10.

$$\frac{dI}{dt} = \beta SI - \frac{I}{r} - qI \tag{4.10}$$

The constant η is incorporated into equation 4.11 to represent the recovery rate of quarantined people.

$$\frac{dQ}{dt} = qI - \eta Q \tag{4.11}$$

The number of recovered individuals changes as shown in equation 4.12.

$$\frac{dR}{dt} = \frac{I}{r} + -\eta Q \tag{4.12}$$

4.1.4 Other epidemic models

Because of the complexity of modeling a disease spread in the actual society, there is a variety of models that try to resemble as much as possible the reality. Some models are not going to be crucial for this document but worth mentioning. These models are:

- *Susceptible-Infected-Susceptible (SIS)*: In this model, there is not immunity for the population. In this case, after being infected, individuals get susceptible of the disease again.
- *Susceptible-Exposed-Infected-Recovered-Susceptible (SEIRS)*: Similar to *SEIR epidemic model*, but after a time t of getting recovered, these individuals lose their immunity and get susceptible again.

These previously mentioned models can be used according to the corresponding disease characteristics, hence in order to select the proper model, it is necessary to know some disease trends.

4.2 Approaches

To support the evaluation of vaccination strategies computational systems use three main approaches to model the effects of their strategies in the control of the disease spread. Analytical deterministic [20] and stochastic [11] approaches are deepened in subsection 4.2.1 and simulation approaches [12] in subsection 4.2.2.

4.2.1 Mathematical programming

Mathematical Programming consists of the use of mathematical models, particularly optimizing models, to assist in making decisions.

Mathematical programming models can be solved with direct or indirect methods. In direct methods, the discretization scheme is used to transform the optimal control problem into a nonlinear programming (NLP) problem which can be solved by an NLP solver. In indirect methods, the solution is derived based on the optimally conditions leading to a two-point-boundary problem.

The definition of epidemic models has been done in two ways, deterministic and stochastic:

- Analytical deterministic approaches allow a closed representation of the spread, but require many assumptions. Moreover, these approaches generate a rough model of the current virus spread, but in most cases are very sensitive to parameters changes.
- Analytical stochastic approaches are similar to a deterministic model, but in these cases, parameters are modeled statistically. These approaches are less affected by changes in the model parameters.

4.2.2 Computer Simulation

In a way to simulate the population interactions, a population graph simulation can be constructed. These approaches incorporate realistic assumptions about the population structure and transmission dynamics of a virus. In this network, each node represents a person, and it is connected to other people following specific relation dynamics.

Simulation requires a lot of computation time and therefore simple strategies are preferred in most cases. This makes it not possible to take into account wide characteristics to divide the population.

In figure 4.2 an example of a graph is shown. Here each node can represent a person and the lines are connected to people that have contact with them.

A variety of works that used these approaches are described in the next section.

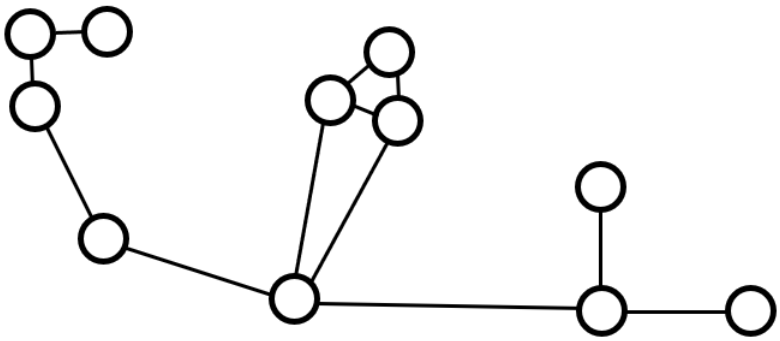


Figure 4.2: Example of graph. Each node corresponds to a person and the edges represent the connections with other people.

4.3 Vaccination models

Once a disease is spreading through the population, it is necessary to stop it. One of the main and classical solutions for stopping/reducing viruses caused diseases along with history, is vaccination.

Vaccination has been demonstrated as a very effective strategy, which prevents the population from being infected [31]. There are different types of vaccines some of them are:

- Live-attenuated vaccines: these use a weakened form of the germ that causes a disease.
- Inactivated vaccines: these use the killed version of a germ.
- Toxoid vaccines: these use a toxin created by the germ that produces the disease.
- Vaccines that use specific pieces of the germ, for example proteins.

The development of any type of vaccine is a long process that involves various stages. These stages include identification of effective antigen, lab-scale antigen, safety concern, animal model study, human trials, or efficacy. This whole process can take at least 18-16 months if it is made lineal, but testing and trials can be performed in parallel to hurry up the process.

Some vaccines require more than one dose, for example, human papillomavirus is administrated in three doses. This, because the reactions of the immune system

to some vaccines are less accurate in getting the immunity, so with those doses the probability of vaccine success increases.

To date, there are about 114 candidate vaccines under pre-crinical stage for SARS-CoV-2, and the development of a 100% effective vaccine can take at least some months or years [2].

In history, some epidemic models have been adapted adding vaccination to determine epidemic control strategies, some of these are discussed bellow.

In 2008 Tanner et al. [26] used a stochastic programming approach for finding vaccination strategies on a SIR Model. They proposed to incorporate stochasticity with the aim of adding robustness to the results obtained.

In this work the notation R_* denotes the threshold parameter for a community of households, post vaccination reproductive number. In their approach, the authors aim to determine a vaccination policy minimizing the probability of occurrence of an epidemic considering resource constraints.

The stochastic programming approach is divided into:

- Vector of random variables $\tilde{\omega}$ with discrete distributions.
- Optimize the random function $z(x, \omega)$
- Stochastic programming solution (SPP): Minimum expected value of $z(x, \omega)$ in terms of $\tilde{\omega}$
- Wait-and-see solution (WS): Expected value of the solutions found, after the random parameters are realized.
- Value of perfect information (VPI): Difference between SPP and WS. It is used to demonstrate how much improvement can be gained if the true values of the parameters are known.
- Expected value solution (EEV): Expected objective value found with the mean point estimations of the random parameters.

In their approach the authors consider three formulations: disease to be controlled, limited vaccination supply and optimization of cost/benefit of the vaccination process. Due to the difficulty of estimating parameters in epidemiological

models and the complexity of human interactions, the authors created a stochastic programming model that allows vaccination strategies to not depend on point estimates of parameter values.

They define the problem using a linear programming formulation and treated some parameters as random variables. The models were solved using CPLEX 9.0 solver. A numerical example was used to show the stochastic effect in the model. In this example, they divided the population by household size, quantity of children, adults and elderly individuals living in the same house.

They determined that it is necessary to take into account parameter uncertainty in order to get closer to reality. As future work they mentioned to test the robustness of their results by performing simulations.

In 2016 an investigation based on social networks relationship was performed by Kim et. al. [17]. The aim of this study was to determine and select who should be vaccinated given a limited number of vaccines. Authors generate a distribution of vaccines over time intervals, introducing a new problem that they denominate the *Social-relation-based vaccine distribution problem (SVDP)*.

The evolution of interactions was projected and reduced to a single directed graph. A *maximum-flow algorithm strategy* was used to evaluate the generated solution. The types of vaccination problems in previous node connectivity models and the differences they have in this new version are also mentioned in this work. First, the fire-fighter problem that *max save* seeks a vaccination strategy over time to maximize the number of uninfected in a given period. Then, the *min budget* seeks a vaccination strategy to save the largest number of members in a subgroup. Secondly, *graph-cut* problems, identify subset of nodes that after being removed, divide the graph in two components. Last, the *data-aware* vaccination problem, consists of distributing a number of vaccines in time in a graph of social networks. This having knowledge of the probability of infection, but in a static graph.

There is a difference with the model mentioned before [26], in this case, the approach focuses on seeking the best distribution strategy for vaccines that are delivered over time, minimizing the number of infected until a complete cure is developed. So far this study is the first algorithm that calculates a distribution plan based on what is currently available and what will be available soon. They also evaluate scenarios that incorporate quarantine and the possibility of getting

infected again (as in SIS Models subsection 4.1.4.

For the simulation of their approach the *stanford network analysis project (SNAP)* was used. A new vaccine strategy with delivery over time, perfect vaccination and quarantine was implemented in their work based on social relationships and predicting the routes of infection and obtaining better results for their SVDP solutions than the random distribution of vaccines.

In 2016 da Cruz et al. [8] propose an evolutionary multi-objective optimization algorithm that uses a stochastic simulation approach for determining vaccination policies for a *SIR model*. In their work the authors minimize the control costs and the number of infected individuals. Some of their main contributions were:

- Incorporate a time invariant sequence and allow different sizes of pulse control in transient state (between the beginning and the steady state) and steady-state phases (states that are established after a certain time in the system).
- Incorporate a stochastic variation which converges, in average, to SIR Model. The model is called Individual-Based Model (IBM), a structure which simulates relationships between individuals and that requires a large amount of data.
- Include a Hash table avoiding the re-evaluation of repeated solutions.

In order to express the system as a fraction of each group, the authors divide the equations of a *SIR model* incorporating birth and death rate in the whole population. And as the initial conditions they considered to susceptible (S), infected (I) and susceptible (S) as $(S_0, I_0, R_0) = (0.99, 0.01, 0)$ in percentage from the total population.

During the transient control state, they vaccinate a large group of individuals in time intervals using different vaccination coverage and quantity of people. The goal is to get the proportion of infected population lower or equal to 0.01 at the end of this state.

The two objectives to be optimized were the time interval between campaigns and the fraction of susceptible individuals to be vaccinated.

Moreover, the optimization process was divided into two states:

1. Determine a set of *non-dominated* steady-state control policies. In multi-objective optimization a *non-dominated* solution is a solution vector for which there is no other solution that simultaneously get better objective values for all the objectives considered.

2. Determine a subset of control policies, based on the control policies obtained previously.

In order to solve the optimization problem, authors split the continuous time interval into stages. During the steady-state they define a fixed time interval and a constant vaccination percentage for each campaign. The goal is that the number of infected individuals tends to zero. They determined 150 time units as the time considered. This time was divided in 100 $t.u.$ for the steady-state and 50 $t.u.$ for the transient state. For this model, the fixed time interval between vaccination campaigns and fixed percentage of susceptible population which must be vaccinated in each campaign.

Authors evaluated the approach for a population of $N = 1,000$ and the parameters shown in their paper [8]. They used 1,000 iterations in a *Monte Carlo* procedure. From the solutions obtained they perform an analysis with the IBM simulation keeping only robust solutions (non-dominated).

From these experiments they show the frequencies of number of vaccinations carried out on time in which the disease is eradicated and the number of vaccinations performed in which the population remains bellow the tolerable value for each state. 23,645 different solutions were evaluated in the first step and 127,980 in the second step. A vaccination rate of 0.8595 and time interval of 7.1430 $t.u.$ was recommended.

The same year Correa et al. [7] studied a vaccination campaign using an optimization method. Some of their main contributions were:

- Control and state constraints can be introduced in the model.
- Multiple pulse vaccinations can be implemented as control strategy.
- The time points where the population is vaccinated of multiple pulse vaccinations can be optimized, and also, allows different vaccination levels.

As the initial condition they used as susceptible (S), infected (I) and recovered (R) population the values $(S_0, I_0, R_0) = (0.9N, 0.1N, 0)$.

They assumed that susceptible individuals who have been vaccinated become immune and turn directly into recovered individuals.

The population at each vaccination time t_k with a Δp_k new vaccinated people, is defined as shown in 4.13.

$$\begin{aligned} S(t_k^+) &= S(t_k^+) - \Delta p_k \\ I(t_k^+) &= I(t_k^+) \\ R(t_k^+) &= R(t_k^+) + \Delta p_k \end{aligned} \tag{4.13}$$

An *impulsive dynamic system* consists of three elements: continuous time dynamic equation administrating the system dynamics, difference equation representing discrete events and how the states changes and discrete time points.

Authors used the *multiple shooting and collocation (CMSC)* method to solve the optimization problem. It uses multiple shooting to represent the states with individual shoots over sub-intervals and adding finite elements to each sub-interval. The overall problem is solved by an *NLP solver* based on the solutions and the sensitivities computed in each sub-interval.

They evaluated their approach considering a *SIR model* with $N = 1,000$ and 10 impulsive vaccinations, with the values of Δt_k and Δp_k bounded between 1 and 10, and 5 and 500 respectively. From their results, the vaccination with time instants reduces the number of infected people in time and can be improved by using variable-time points of vaccination. Moreover, at the end the number of infected individuals does not stop implying that is necessary to keep the vaccination to control the disease spread. Experiments considering 25 impulsive vaccinations and including the constraint that at the final time the infected population must be zero allow then to conclude that the proposed impulsive dynamic model allows an efficient treatment for transient and steady state. Moreover, this can be used to set health policy actions in disease campaigns.

Another work on 2016 by Wu et. al [30] used a relational network graph (scale-free network). Immunization strategies were proposed for the *SIR model* environment and *SIS epidemic model*. This because there are viruses that can infect

people again after a certain time. Static immunization occurs before the outbreak and dynamic immunization during the outbreak. Different cases of immunization schemes were evaluated through direct immunization schemes in networks.

An *SVIR model* was formulated (vaccination is added to SIR model). Here, they represent immunized population, and apply random vaccination along with targeted vaccination strategies. A path parameter was incorporated to be aware of immunizing high-risk nodes. This is the effective method in scale-free networks to vaccinate the nodes with the highest degree, the one connected to more nodes, as can be seen in figure 4.3. This group can be defined as a region or group of people sharing some characteristics. Then, the set of nodes that must be vaccinated in a graph to decrease or remove the virus are identified.

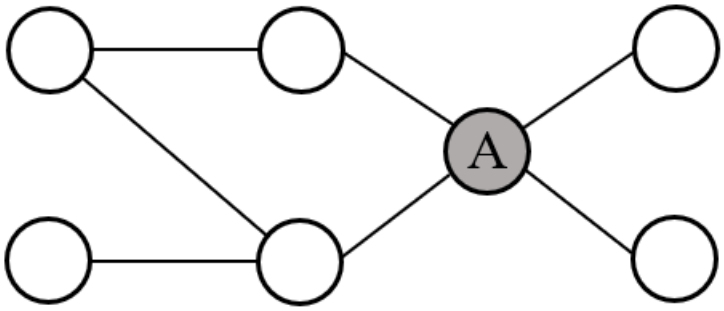


Figure 4.3: Graph example, where node A is the node with highest degree.

In their approach Wu et. al create and test their model with Monte Carlo simulations on a population of $N = 2,000$ in a scale free network from a standard configuration model with a degree exponent of 2.7.

They define a simple and deterministic *SIR epidemic network model* and formulated a real-time immunization model, where specific nodes of a subgroup of the population can be vaccinated due to real life situations with limited vaccine quantity and conclude that the local immunization program can highly improve the efficiency of static immunization for a small quantity of vaccines.

In 2018 Wang et al. [29] developed a model for finding optimal vaccination strategies of a constrained time-varying *SEIR epidemic model*. In this model the population (N) is not necessary constant as in some previous works, hence it is defined as $N(t) = S(t) + E(t) + I(t) + R(t)$.

They assumed death rate, birth rate, disease death rate, incubation period and recovery rate as constant values. The vaccination rate is considered the control variable. The contact rate can vary seasonally as a cosine function. Vaccination also changes. Vaccinated individuals get immune with a certain ratio that decrease in time t, because of drug resistance as a monotonic decreasing function. They define limited supply of vaccines at each time instant. As a constraint for keeping the infected individuals at a low level, they used $S(t) < S_{max}$, hence the susceptible population needs to be vaccinated and be lower than S_{max}.

They extended the SEIR model adding time-varying between seasons, with their numerical simulations they prove that omitting time-varying factor may result with a unreasonable vaccination strategy.

Also in 2018 Ng et. al. [23] developed a multiple criteria mathematical programming model to find the optimal combination of influenza vaccination strategies using a deterministic approach. As performance measures they considered:

- *Program Vaccination Cost*: Minimizing the program cost.
- *Vaccination Efficacy (VE)*: Fraction of people no longer susceptible and immune due to vaccination. To estimate the VE they used the post-vaccine reproduction number [16], where when it is lower, there is a lower morbidity and the disease stops faster.
- *Societal Benefits (SB)*: Number of prevented visits to doctors and the mean number saved working days.

Thus, they formulate a multi-criterion optimization problem and search for non-dominated solutions focusing on determining the optimal number of vaccine doses to be assigned to different population groups at risk. As it was mentioned before, a solution that can not be improved in a criterion without worsening at least another one is considered a non dominated one.

Furthermore, they define three types of vaccination strategies:

- *Mass vaccination*: Vaccinate the whole population over a short period of time.
- *Random vaccination:* Vaccinate randomly members of the community. It should be possible to vaccinate isolated people or people that not belong to risk groups.

- *Targeted vaccination:* Vaccinate a small group of people that can impact in the spread of the virus or can be more affected by the disease.

The planning is studied for the three main stages of an influenza season, beginning, peak and ending stages, between October and March corresponding to Israel.

They focus on three groups of vaccination strategies, the first group corresponds to evaluate vaccine composition strategies looking for the most effective, where the release date of the vaccine and its efficiency should be taken into account. The second group studies strategic planning based on socio-economic analysis of vaccine supply chain. Finally, the third group and main criterion is to identify the optimal vaccination strategy exploiting a graph representation of the disease spread.

In their work they use a deterministic linear programming optimization rather than simulation owing to simplify the problem without using complex networks models and lower computational resources.

Finally, authors establish that successful immunization strategies can be developed only by taking into account the inhomogeneous connectivity properties of scale-free networks. Also, school-age children have highest transmission rates and vaccinating elders first gives lowest impact of transmission. Furthermore, between mortally-based and targeted at high-risk strategies, mortally-based is better for a highly transmissible disease. Moreover, some previous articles mention strategies within age-structured populations and time together with quantity of vaccine doses available.

The same year, Chen et. al. [4] introduced an epidemiological model that considers a variable population size, imperfect vaccination respect to the degree of a node and quarantine in scale-free networks for models with and without permanent immunity. These models are called the *Susceptible-Vaccinated-Infected-Quarantined-Recovered (SVIQR)* and *Susceptible-Vaccinated-Infected-Quarantined (SVIQS)*.

For the *SVIQR* model four factors were considered: demographic impact, epidemic framework, vaccination and quarantine. Depending on the virus, the demographic impact may have a higher or lower influence as it is the case with HIV but not with influenza.

The model corresponds to the framework of a *SIQR* model and incorporates their same states. The vaccination factor considers that the vaccination will be imperfect, obtaining a degree of immunity with a certain probability. Moreover, in each new period the infected will be quarantined with a certain probability. For the *SVIQS* model the same factors are used, but the epidemic framework is changed to *SIS*. Theoretical results for the proposed mathematical models were generated and a stochastic Monte Carlo simulation was performed to complement the results obtained.

They showed that R_0 is related to the topology of networks after the simulations. Also, quarantine and vaccination rate decrease R_0, but quarantine produces a higher impact on decreasing the R_0 number in comparison with the implemented vaccination policies.

In 2019 Enayati et. al. [10] modified the *SEIR model*, incorporating a quarantine factor.

In this work, the population is separated into subgroups. Each subgroup has its own proportion of susceptible (S), exposed (E), infected (I), quarantined (Q) and recovered (R) individuals. Also, they change the generic contact rate of the *SEIR model* to define a more specific contact rate between subgroups, with these values they also determine the rate of new infections in each subgroup.

This model seeks the optimal vaccination distribution plan for subgroups differentiated by geographic location and age, in order to extinguish the influenza outbreak (They used R_0 and other parameters of influenza).

In their study, authors considered the number of new infections generated by an infectious individual in order to evaluate vaccine distribution strategies. For this, it is calculated the next generation matrix, that determines the number of secondary infections in consecutive generations due to the interactions of each subgroup with the other subgroups.

To obtain the solution a nonlinear mathematical model is used, minimizing the total number of doses. These solutions do not guarantee to be the optimal but they are not far from the global optimal solution for instances with three to six regions.

Also, the authors developed a global optimization algorithm based on discretization with multi-parametric dissagregation. They evaluate the different vaccination policies using their implementation.

The method used can be seen as a recommender system for public health authorities of the number of vaccines needed to stop future influenza outbreaks and it is enough flexible to change some subgroup characteristics and use it in real life policy decisions. As future work, authors consider to include uncertainty (stochastic parameters) such as vaccine efficacy and also apply the model to other infectious diseases.

4.3.1 COVID-19

Because of SARS-CoV-2 virus spread causing COVID-19 disease is a topic of concern, some epidemic models have been studied.

Contreras et. al. [6] proposed a new variation of the *SEIR model* with asymptomatic individuals (A) called *SEIRA epidemic model*.

To model asymptomatic population a constant α divides the infected individuals as $I(t) = (1-\alpha)I + \alpha I$. Where the left side corresponds to symptomatic infected individuals and the left side corresponds to asymptomatic individuals. In their study they also use specific parameters for modeling SARS-CoV-2 spread.

Their model divide the population into subgroups related by different factors as geographical or economical. They define particular interaction factor between each subgroup. For this an interaction matrix is calculated.

As an experiment to show how the model works, they tested a small example for three subgroups: in normal circumstances, isolating one of the subgroups and quarantine.

The proposed model can represent interactions between varied and new subgroups and also it can be used for other diseases using proper parameter values. As future work, behavioral changes can be added through new parameters.

Also, Mishra et. al [21] proposed another variation of *SEIR model* including:

- Home isolation (Q_1): Susceptible people or possibly exposed to the disease staying at home.
- Quarantine (Q_2): Infected people who is set in quarantine.

They developed three epidemic models dividing the infection of COVID-19 in three different stages:

1. New infections from people who travelled to other countries and comes back to the country.
2. Local transmission due to people who get in immediate contact with infected persons the first stage.
3. Community transmission due to people who has not travel history and also not contact with people of the first stage.

They executed simulations for the three stages considering R_0 values bellow and over 1. As a result, when $R_0 > 1$ an infection-endemic equilibrium was found for all stages. Also, it was demonstrated that a higher home isolation and quarantine implies lower chances of disease transmission.

4.4 Chapter Summary

In this chapter we described some of the most used types of epidemic models, variations of SIR Model and other types. Also, approaches for determining vaccination strategies were described, classified into two main categories: mathematical programming and computer simulation.

After understanding the previous points, the state of the art of vaccination models was presented. Table 4.1 shows a summary of the main features of approaches here described. The first column shows the reference of the paper, the second row the year it was published. The third row shows the kind of efficacy the vaccine has in the model. Following, the vaccination type if it considers time periods or not. Fifth column shows the type of epidemic model considered. After this, is the type of approach: mathematical or computer-based approach. The seventh column shows the type of subgroups if there are. Finally, the optimization goals are listed.

It can be noted that most of these works considered perfect vaccination, i.e., vaccination that works for 100% of the population leaving them into immune state.

Further, most of them determine a vaccination rate for the population instead of dividing the quantity over time.

Different epidemic models have been considered when applying vaccination strategies, but in the latest works SEIR models seem to be more used for resembling the reality. Moreover, quarantine has been analyzed in different works proving to affect notoriously in the spread of the disease.

Some works that used analytical deterministic, analytical stochastic and computer simulation were studied. The use of these different approaches limits the number of the population subgroups, due to computer complexity of computer simulations.

Some works divide the population into subgroups, the most used features were the age and geographical regions. When the age is considered, it can be determined death rate for some specific diseases. Moreover, regions are used because of the possibility of limiting interactions between them.

As objective, most of the works considered the search for a vaccination plans that minimize infected population or the total infected population along time. For some cases, the minimization of the number of vaccines is wanted, but most works consider a maximum quantity of vaccines available as constraint. It is also possible to find solutions for more than one objective in multi-objective approaches that, for example, minimize the infected population and maximize the societal benefits [23].

Finally, two epidemic models that were created for modeling COVID-19 spread are mentioned. Both use the SEIR model with different adaptations as quarantine, asymptomatic people and external people that get into a country.

Reference	Year	Vaccination Efficacy	Vaccination Type	Epidemic Model	Approach	Subgroups	Objective Function
Tanner et al. [26]	2008	Efficacy based on normal distribution with mean 0.85	Vaccination rate	SIR model	Stochastic programming	Household size, age	Minimize the probability that an epidemic will occur
Kim et al. [17]	2016	Perfect vaccination	Groups over time	SIQR model SIQS Model	Population graph Simulation	Sub graphs in temporal graph	Minimize infected people
Da cruz et al. [8]	2016	Perfect vaccination	Groups over time	SIR model	Stochastic simulation	-	Minimize vaccination cost campaigns and the integral of infected population
Correa et al. [7]	2016	Perfect vaccination	Groups over time	SIR model	Optimal control problem solve NLP	-	Minimize the integral of infected population
Wu et al. [30]	2016	Perfect vaccination	Vaccination rate	SIR model SIS model	Population graph with random and targeted immunization	Create set of nodes	Minimize total number of vaccines Minimize the total cost
Wang et al. [29]	2018	Exponential vaccine resistance	Vaccination rate	SEIR model	Nonlinear constrained optimal control	-	Minimize the integral of infected population
Ng et al. [23]	2018	Perfect vaccination	Vaccination rate	SIR model	Multiple criteria mathematical programming model	Age	Maximize societal benefits Minimize costs Maximize vaccination efficacy
Chen et al. [4]	2018	Perfect vaccination	Vaccination rate	SIQR model SIQS Model	Network simulation	-	Minimize amount of infected population
Enayati et al. [10]	2019	Perfect vaccination	Vaccination rate	SEQIR model	Mathematical program	Regions, Age	Minimize amount of vaccines used

Table 4.1: Summary table of vaccination models presented in Section 3.3.

Problem Definition

This chapter describes the problem to be solved as a susceptible (S), exposed (E), infected (I), asymptomatic (A) and recovered (R) model with subgroups, vaccination and restrictive measures that change according to the quantity of infected population and named as SEAIRV model.

First, the behaviors the model follows are explained and related with previous models and parameters of Chapter 4. After it, how the interactions between subgroups are modeled. Posterior to it, the mathematical model is deepened, obligatory and optional parameters are introduced, also, variables and constraints. Finally, the objective function of the mathematical model is shown.

5.1 Explanation

For modeling COVID-19 spread a Susceptible-Exposed-Infected-Recovered model (SEIR Model), introduced in Chapter 4, with asymptomatic population and variable quarantine is implemented. The model follows the next obligatory rules:

- Birth rate is equal to death rate.
- Exposed individuals split into infected and asymptomatic individuals with their perspectives rates.
- A fraction of asymptomatic individuals get tested and move to Infected individuals.
- There is a limited quantity of vaccines.

- Population is divided into subgroups. This division can be based on geographical zones, political divisions like cities or districts.
- Each zone has its own contact rate. It has a contact rate with itself and with all other zones. These contact rates can be interpreted as the percentage of people in the group that stays in the zone and the percentage of people in the group that visit, travel or work into another zone.
- Each country has restricted in greater or lesser way the movement of the population. Following the methodology Chile has used in which each district has more restrictions based on the amount of infected population they have. For this, population in each subgroup decreases their contact rate according to their infection level. If the zone has more infected population the restrictions should make the population contact decrease. In this work, for each zone will be used a mathematical equations 5.13 and 5.14, that resembles this behavior, deepened later in Section 5.3.
- Infected population decreases its movement considering an equal or higher rate compared to infected population.

Additionally, vaccination in periods of time is implemented, the algorithm per defect makes just one vaccination at the beginning of the epidemic. It searches only for the best vaccination plan with a fixed quantity of vaccines to assign to different subgroups every period.

Furthermore, some of the rules are optionally editable and available in the model:

- Stochasticity: add an stochasticity factor to the infectiousness of the virus.
- Rigor of protection measures: change the parameters of how restrictive are the measures for infected population and the not infected population. These parameters are the maximum percentage of the infected population that triggers the movement restriction, the values MIP_α of the movement of non infected and MIP_β of infected in the mathematical equations 5.13, 5.14 explained in section 5.3.
- Vacation period: add periods where the health protection measures decrease owing to vacations. In the vacation period the people increases the movement rates between zones and increases their interaction.

The model phases of the population are shown in figure 5.1. Population can be susceptible (S), exposed (E), asymptomatic (A), infected (I) or recovered (R). Susceptible individuals get into exposed individuals because of the interaction with infectious population. The infectious population can be considered infected individuals with infection rate β and asymptomatic individuals with infection rate α. Exposed individuals get infectious with a rate of δ. A μ part of the infectious population are asymptomatic and $1 - \mu$ show symptoms. From the asymptomatic population, some of them get tested with rate η and start figuring as infected. Infectious population get recovered with rate γ.

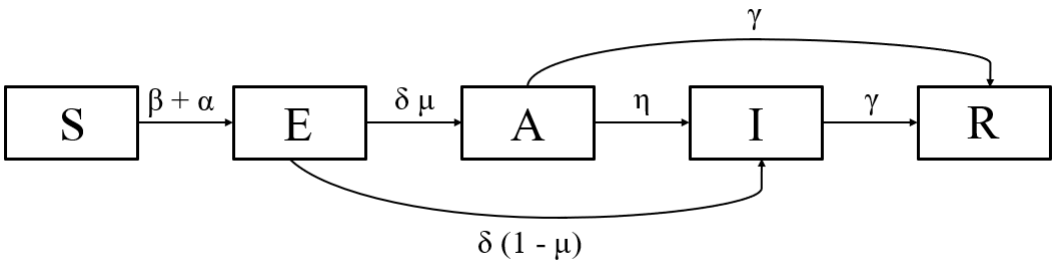


Figure 5.1: SEAIR model. S, E, A, I, R corresponds to susceptible, exposed, asymptomatic, infected and recovered population respectively.

5.2 Interactions

In order to model the interactions between subgroups in population, the next assumptions were assumed.

- People from each subgroup can visit any other subgroup. This visits can be due to work places mainly.
- The percentage of people that move to another subgroup is represented on a contact matrix, which sums at the beginning 100% per each subgroup.
- For any subgroup x, subgroup x' corresponds to the group formed by all the population that visits subgroup x excluding the population from subgroup x that visits another subgroups.
- The newly infected population of subgroup x' is obtained from the contact between infected and susceptible people of subgroup x'.

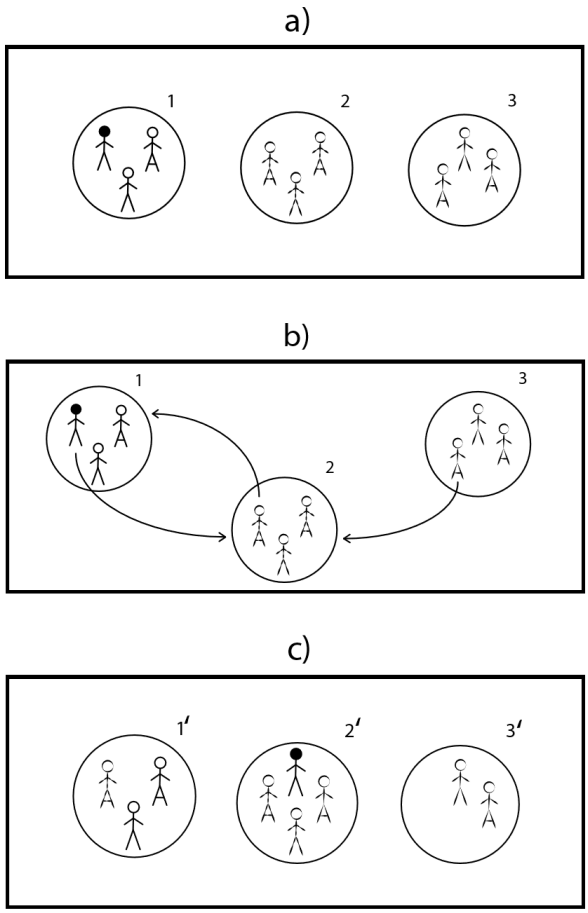


Figure 5.2: Example of a interaction between three subgroups.

- The newly infected population of subgroup x is computed as the sum of all new infected from subgroup x that visits other subgroups and the new infections of people that stay at subgroup x.

Figure 5.2 shows an illustrative example of these interactions. Figure 5.2 a) shows three different subgroups and subgroup 1 has one infected individual. Figure 5.2 b) shows one person from subgroup 1 and one person from subgroup 3 visiting subgroup 2. Moreover, there is also one person from subgroup 2 visiting subgroup 1. The new subgroups obtained from these interactions (subgroup 1', subgroup 2' and subgroup 3') are then shown in figure 5.2 c).

5.3 Mathematical model

To manage the rules previously explained, it is necessary to fulfill the rules described by the following mathematical model. First, we describe the parameters defined by the problem. The decision variables related to the

5.3.1 Parameters

- N: Population size.
- G: Set of subgroups.
- V: Quantity of vaccines available per period.
- T: Maximum time lapse of the model simulation.
- ζ_{ij}: Percentage of people from subgroup i that go to subgroup j, that conforms the contact matrix.
- β_{ij}: Infectious rate between subgroup i and j.
- α_{ij}: Infectious asymptomatic rate between subgroup i and j.
- δ: Incubation rate, rate based on days of the disease incubation on a exposed individual to get infectious.
- γ: Recovery rate, rate based on days of an infectious individual (infected or asymptomatic) to get recovered.
- η: Detection of asymptomatic rate. The detection based on how many tests are performed to asymptomatic individuals, also depends on the real number of asymptomatic individuals.
- μ: Percentage of infectious individuals that are asymptomatic.
- MIP_{α}: Percentage of the maximum infected people of the subgroup that triggers movement restrictions of non infected population of the subgroup (i.e quarantine).
- MIP_{β}: Percentage of the maximum infected people of the subgroup that triggers movement restrictions of infected population of the subgroup (i.e

quarantine). More restrictive measures should be implemented for infected people.

- $G, P, V, T \in \mathbb{N}$
- $\beta_{ij},\ \alpha_{ij} \in \{0,1\},\ \forall\ i,j \in G$
- $\delta,\ \gamma,\ \mu, \in \{0,1\}$

Optional parameters:

- P: Number of vaccination periods, each period can get different distribution of the vaccines per subgroup and produce important changes on the solutions. Also the vaccines are being elaborated and the arrival of vaccines can depend of production chains.
- B: Duration of the vacation period when protection measures will be decreased in days.
- ω: How strong protection measures are going to be decreased on vacation period. Indicate how much the movement of the population will increase during vacations period.
- S: Stochasticity.

5.3.2 Variables

- v_{ip}: Number of persons from subgroup i that will be vaccinated in period p.

5.3.3 Constraints

Classical SEIR model constraints

- The population should maintain their size for this reason the quantity of susceptible (S), exposed (E), asymptomatic (A), infected (I) and recovered people sum the total quantity of the population at all times.

$$N = S(t) + E(t) + A(t) + I(t) + R(t) \tag{5.1}$$

- Susceptible population of each subgroup gets infected depending on their contact rate with infected in other subgroups and themselves.

$$\frac{dS_i}{dt} = -\sum_{j=0}^{G}\left(\frac{\zeta_{ij}(t)\ S_i(t)}{\sum\limits_{k=0}^{G}\Big(\ N_k\ \zeta_{kj}(t)\Big)} \sum_{k=0}^{G}\Big(\ \beta_{kj}(t)\ I_k(t)\ +\ \alpha_{kj}(t)\ A_k(t)\Big)\right) \tag{5.2}$$

- Exposed individuals per each subgroup get infected or asymptomatic after the incubation period.

$$\frac{dE_i}{dt} = \frac{dS_i}{dt} -\ \delta\ E_i \tag{5.3}$$

- A fraction of exposed individuals of each subgroup gets into asymptomatic individuals.

$$\frac{dA_i}{dt} = \mu\ \delta\ E_i\ -\ \gamma\ A_i - \eta\ A_i \tag{5.4}$$

- A fraction of Exposed individuals of each subgroup is tested as positive COVID-19 cases and gets into infected individuals.

$$\frac{dI_i}{dt} = (1\ -\ \mu)\ \delta\ E_i\ +\ \eta\ A_i -\ \gamma\ I_i \tag{5.5}$$

- Infected and asymptomatic individuals per subgroup get recovered after the recovery period.

$$\frac{dR_i}{dt} = \gamma\ (I_i\ +\ A_i) \tag{5.6}$$

Contact rate constraints

- People that move from subgroup i to subgroup j decrease because of government regulations that depend on the quantity of infected.

$$\zeta_{ij}(t) = \zeta_{ij}(0)\ F_{\alpha i}(I_i(t)) \tag{5.7}$$

- The infection rate between subgroups (Beta) is obtained from the R_0 value, the recovery rate and the contact between them.

$$\beta_{ij}(0) = R_0\ \gamma\ \zeta_{ij}(0) \tag{5.8}$$

- Contact rate of infected people decreases due to government regulations that

depending on the number of infected people.

$$\beta_{ij}(t) = \beta_{ij}(0)\ F_{\beta i}(I_i(t)) \tag{5.9}$$

- Contact rate of asymptomatic people decreases because of government regulations that depend on the number of infected people.

$$\alpha_{ij}(t) = \beta_{ij}(0)\ F_{\alpha i}(I_i(t)) \tag{5.10}$$

- Contact rate of infected is lower or equal to the asymptomatic people.

$$\alpha_{ij}(t) \geq \beta_{ij}(t) \tag{5.11}$$

- When there are not infected people or are really low, the contact keeps normal.

$$F_{\alpha i}(I_i(t)) \approx 1\ ;\ I_i(t) << N \tag{5.12}$$

- Government restrictions can be triggered when there is a high number of infected population. Anyway, either for mental health reasons and work responsibilities, people increases their movement when the infected population decrease.

$$F_{\alpha i}(I_i(t)) = \frac{100}{100 + A_1^{-B_1 \frac{I_i(t)}{MIP_\alpha\ N_i}}} \tag{5.13}$$

$$F_{\beta i}(I_i(t)) = \frac{100}{100 + A_2^{-B_2 \frac{I_i(t)}{MIP_\beta\ N_i}}} \tag{5.14}$$

Vaccination constraints

- At the beginning of each vaccination period, the vaccinated people get immediately recovered.

$$S_i(t_p) = S_i(t_{p-1}) + \frac{dS_i}{dt} - v_{ip} \tag{5.15}$$

$$R_i(t_p) = R_i(t_{p-1}) + \frac{dR_i}{dt} + v_{ip} \tag{5.16}$$

Optional

- Adding stochasticity to the new infected people.

$$\frac{dS_i}{dt} = \frac{dS_i}{dt}\ random(0.9-1) \tag{5.17}$$

5.3.4 Objective Function

In this work, two objective functions can be considered.

- Minimize the quantity of infectious people in all the periods.

$$Obj_P = min\ \{I(t_f) - A(t_f)\} \tag{5.18}$$

- Minimize the quantity of infectious people at the same time.

$$Obj_S = min\ \{max_t\{I(t) + A(t)\}\} \tag{5.19}$$

5.4 Chapter Summary

In this chapter we define the problem and named it as SEAIRV model. The model needs considers rules like dividing the exposed people into infected and asymptomatic and having a limited number of vaccines. Also there are some optionally rules like vacation periods. Besides these rules, the model follows transitions between each of the population types (susceptible, exposed, infected, asymptomatic, recovered) through rates. In a way of applying interaction between subgroups the model is formulated considering the contact between subgroups and all the individuals that interact in each subgroup.

The decision variable defines the number of people from each subgroup that get vaccinated on each period (v_ip). Additionally, the contact matrix is used considering the incubation rate (δ) that determines the number of exposed individuals to get infected or asymptomatic. The recovery rate (γ) determines the quantity of infected and asymptomatic population that gets recovered. Also, the reproduction number (R_0) value along with the contact of subgroups (ζ_{ij}) controls the infectious rate (β). Furthermore, the functions $F_{\alpha i}$ and $F_{\beta i}$ were introduced to resemble how

the restriction of the movement measures are increased or decreased depending on the infected population. $F_{\alpha i}$ is associated to the restrictions to non infected population and $F_{\beta i}$ those associated to infected population. Finally, two objective functions were considered, the minimization of the number of infectious people and the minimization of the number of infectious people at the same time.

Algorithm Description

This chapter describes the algorithm used to get an efficient vaccination plan specifying initially the number of vaccines and periods. Through a tabu search implementation we seek for a solution able to decrease effectively the spread of SARS-CoV-2 virus.

First, we explain the representation of solutions. Secondly, the evaluation function that considers two main objective functions. Third, we deepen on the structure of the approach including each a detailed explanation of all the processes implemented.

6.1 Representation

Solutions are represented as arrays of vaccination distributions. The size of each solution depends on both, the number of subgroups (G) and the number of periods (P) previously explained in Section 5.3. Each cell indicates the number of vaccines allocated to each subgroup in each period. Figure 6.1 shows an example of solution for a problem that considers five subgroups, two periods and 100 vaccines per period. Here, for each period, an array of vaccination distribution represented as an array determines the of number of vaccines per subgroup. A feasible solution keep the same total of vaccines in each vaccination period.

Solutions are represented as arrays of vaccination distributions. The size of each solution depends on both, the number of subgroups and the number of periods defined for each problem instance. Here each cell indicates the number of vaccines allocated to each subgroup in each period. Figure 6.1 shows an example of solution for a problem that considers five subgroups, two periods and 100 vaccines per

period. Here, for each period, an array of vaccination distribution represented as an array determines the of number of vaccines per subgroup. A feasible solution keep the same total of vaccines in each vaccination period.

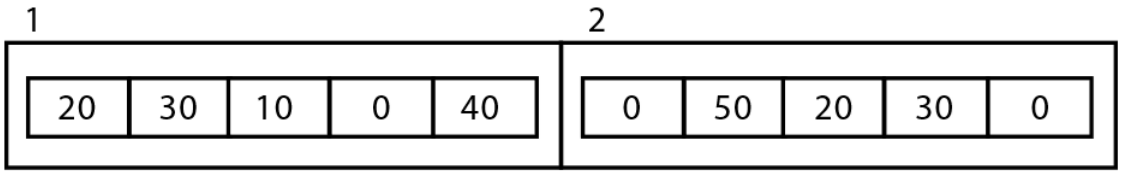


Figure 6.1: Example of a solution representation for a problem with five subgroups, two periods and 100 vaccines per period.

The solution in figure 6.1 allocates in the first period 20 vaccines to *subgroup* 1, 30 to *subgroup* 2, 10 to *subgroup* 3 and 40 to *subgroup* 5. In the second period it allocates 50 vaccines to *subgroup* 2, 20 to *subgroup* 3 and 30 to *subgroup* 4. Moreover, in both periods, the total allocated vaccines are 100.

6.2 Evaluation Function

In this work we consider two objective functions: minimize the quantity of infected population at the same time (Obj_P) and minimize the quantity of infected population in all the period (Obj_S). These objective functions are shown in section 5.3.4 of chapter 5. From this, we define the evaluation function as a weighted sum of the these both objective as shown in equation (6.1).

$$Obj_P\ Factor1 + Obj_S\ Factor2 \tag{6.1}$$

The number of maximum infected at the same time and total infected in the period are computed by the problem model explained in section 5.3.

6.3 Algorithm structure

The approach proposed to face the vaccination allocation problem is a local search based approach that separates each problem according to the number of periods

it considers. As shown in algorithm 1, for each period a set of initialization, local search and completion steps are performed. The whole process starts defining the initial values of the period, the quantity of susceptible, exposed, infected, asymptomatic and recovered population at the beginning of the current period (line 1). Next, an initial solution is constructed with the initialization method explained in section 6.3.1 (line 5). A local search process it then performed on the constructed solution. This method returns the best solution found during its search. This method is described in section 6.4 (line 6). The previously best solution found is set as the vaccination for that period (line 7) and used to get the initial values of the next period (line 8). The previous process is performed for each period in the problem instance. The final result contains the complete vaccination plan for all periods.

Algorithm 1: SEAIRV Search

Input : Population, R_0, Gamma, Delta, Eta, ContactMatrix
Output: Vaccination distribution for each period

```
1 nextInit ← setInitialvalues(input)
2 completeSolution ← Empty list
3 foreach Period do
4 |   init ← nextInit
5 |   sol ← initializeSolution(Method, init)
6 |   sol ← LocalSearch(sol, iterations, Stopcriteria)
7 |   completeSolution ← Add(sol)
8 |   nextInit ← setInitialvalues(sol)
9 end
```

6.3.1 Initialization

To initialize the solution, five methods are proposed. These methods are studied in order to find, if exists, an initialization process that produce better solutions in lower time and iterations. This, because of the high computational times required by the evaluation function when solving large size problem instances.

Initialization methods proposed are listed below:

- Inner Interaction: This method allocates vaccines to each subgroup proportionally to the interaction between other subgroups and the current one. It

starts computing the sum of all the interactions from other subgroups to the current subgroup. Then, it calculate the percentage according to the value of inner interaction per subgroup. Then it allocates vaccines with probabilities based on their percentage.

- Outer Interaction: This method allocates vaccines to each subgroup proportionally to its interaction with the other subgroups in the problem instance. It starts computing the sum of all the interactions from the current subgroup to other subgroups. Then, it computes the percentage according to the value of outer interaction per subgroup. at last it allocates vaccines with probabilities based on their percentage.

- Mixed Interaction: This method allocates vaccines to each subgroup proportionally to the other groups interaction with the current one and its interaction with the other subgroups in the problem instance. It starts computing the sum of all the interactions from the actual subgroup to other subgroups and other subgroups to the actual subgroup. Then, it computes the percentage according to the value of outer interaction per subgroup. It allocates vaccines with probabilities based on their percentage.

- Equity: This method allocates vaccines an equitable probability for the number of vaccines to all the subgroups in the problem instance. For this, we give the same probability to get a vaccine to each subgroup. Then, it allocates vaccines following a uniform distribution.

- All to One: This method allocates all the vaccines of the period to just one subgroup. For this, it evaluates all the subgroups and select the option that obtains the best evaluation function.

6.4 Local search

The local search process is shown in the algorithm 2. The local search procedure is based on the tabu search metaheuristic. At the beginning of the process, an empty tabu list is considered (line 3). The search considers as stopping criteria a maximum number of iterations and a stuck criteria (line 4). This last criterion is evaluated as the number of iterations without changing the current best solution. At each iteration a new solution is generated (line 6) using one of the possible movements generated with a random function with probabilities for each (line 5).

A set of three movements were proposed in this approach. These movements are explained in section 6.5. The applied movement is recorded into the tabu list (line 6). The tabu list records a pair formed by the subgroups that trade vaccines, i.e. the pair of subgroups involved in the movement. For example, if a movement was generated between subgroup 2 and subgroup 6 , without regarding the movement type, it records the tuple (6,2) because subgroup 6 can not return vaccines to group 2. Neither a swap nor an invert can be performed between them.

Moreover, if the obtained solution is better than the actual best, the best solution is updated. In other case, the number of iterations without improvement is increased (line 7). At the end of the process it returns the best solution found.

Algorithm 2: Local Search

```
   Input  : sol, Iterations, Stopcriteria
   Output: Vaccination distribution for one period
 1 iter ← 0
 2 withoutChange ← 0
 3 tabuList ← Empty list
 4 while Iterations > iter AND Stopcriteria > withoutChange do
 5 |   movType ← randomMovType()
 6 |   sol, tabuList ← generateMovement(movType, tabuList, Objective)
 7 |   withoutChange ← withoutChange + 1
 8 |   if bestValue > Evaluation function (sol) then
 9 |   |   bestValue ← value(sol)
10 |   |   bestVaccination ← sol
11 |   |   nextInit ← initialValues(sol)
12 |   |   withoutChange ← 0
13 |   end
14 end
```

6.5 Movements

In our approach we propose and evaluate three different movements to modify solutions. These movements are explained below:

6.5.1 Give Random

The idea of this movement is to reallocate the number of vaccines between subgroups. For this one random subgroup (sg) having at least one vaccine allocated is selected. Also, a random quantity between 1 and the number of vaccines (v) currently allocated to (sg) is chosen. These v vaccines are reallocated from sg to each other subgroup. Each possible change is considered a neighbor solution. The first new allocation that shows a better evaluation function is selected considering a *first improvement* approach. Moreover, if there is not possible to get any improvement, a *best improvement* approach is considered selecting the best option.

An example of the application of this movement is shown in figure 6.2, for a period of time with four subgroups and 50 vaccines. Here if the subgroup 2 is selected, the neighborhood generated has a size of three (four subgroups minus the selected subgroup). Moreover, the movement evaluates to reallocate 20 of the 30 vaccines of subgroup 2 to each of the other subgroups.

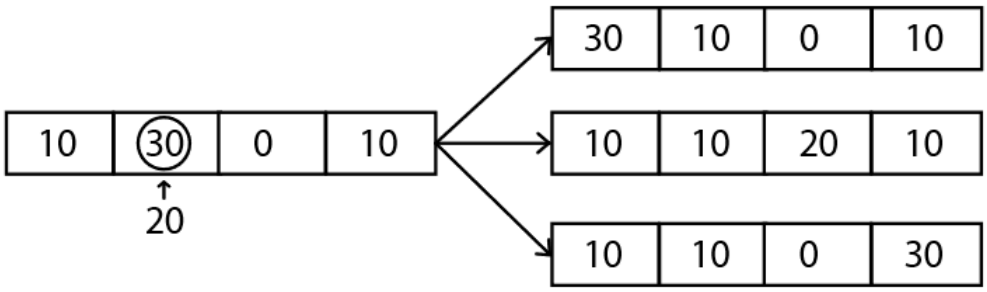


Figure 6.2: Example of give random movement between *subgroup 2* and the neighborhood obtained.

6.5.2 Swap Random

The idea of this movement is to diversify the search and have chances of a faster converge by finding new solutions in the case of being stuck into a local optima. For this, two random subgroups with different quantity of vaccines are selected and they swap their vaccine quantity. In this case the neighborhood is constructed by applying five times this movement, every time for two random subgroups.

For example, in figure 6.3, for a given period with four subgroups and 50 vaccines. If the subgroup 2 and 3 are selected, they swap their vaccines obtaining, in this case, 30 vaccines from subgroup 2 to subgroup 3 and 0 vaccines from sub-

group 3 to subgroup 2. This process needs to be executed five times to construct a neighborhood.

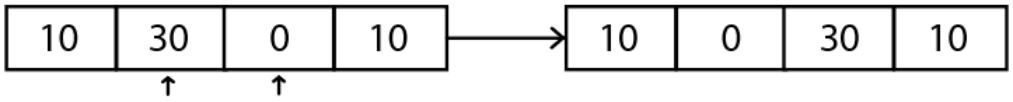


Figure 6.3: Example of swap random movement between *subgroup 2* and *subgroup 3*.

6.5.3 Invert between two Random

The idea of this movement, as the one mentioned before, is to diversify the search and allow a faster converge by finding new solutions in the case of being stuck into a local optima. For this, two random subgroups are selected and they generate an inversion of the number of vaccines to all the subgroups between. This works similar to a mirror effect and the neighborhood is constructed again by applying five times the movement. For example, in figure 6.4, for a period with five subgroups and 75 vaccines, if the subgroup 1 and 4 are selected, they swap their vaccines but also the subgroups between them resembling a mirror, obtaining in this case the changes between subgroup 1 and 4, but also, between 2 and 3.

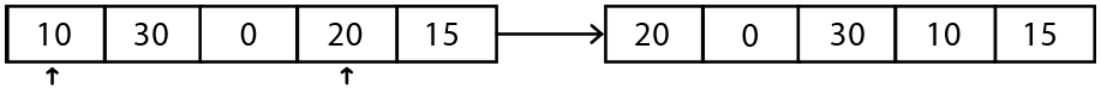


Figure 6.4: Example of invert between movement considering range between *subgroup 1* and *subgroup 4*.

6.6 Chapter Summary

In this chapter we introduced and described the local search algorithms implemented for finding an efficient vaccination plan considering a fixed number of periods, subgroups and vaccines. Solutions were defined as arrays of vaccination distribution where the index is the period and contains an array of vaccines per subgroup. The evaluation function was established as the weighted sum of two

main objectives, minimizing the quantity of infected population in all the periods and minimizing the quantity of infected population at the same time. Five initialization procedures were introduced, based on inner interaction, outer interaction, mixed interaction, equity and all to one strategies. The local search process incorporates three movements that are performed according to fixed probabilities. We proposed a give random, swap random and invert between two random allocations. The approach uses a tabu list to allow diversification accepting low quality solutions avoiding cycles during the search process.

Experiments

In this chapter we present the experimental environment defined to test our proposal. Eight problem instances are used to evaluate our approach and half of them are used to set the SEAIRV parameters. These problems instances were constructed based on real data related to COVID-19 spread during 2020. Here we explain how the input files are composed, we indicate from where the data used was obtained, we explain how the problem instances were constructed and the adjustments performed in COVID-19 and SEAIR model parameters based on real data from COVID-19 test for each input file. Moreover, we also introduce the experiments performed in order to evaluate the main features of the proposed approach.

7.1 Problem instances

7.1.1 File format

As input, a text file with the format in Figure 7.1 is used. The first number corresponds to the number of vaccination periods, the second line indicates the time lapse in days that will be considered each period of the procedure. After it, the number of vaccines to distribute in each period. The number of subgroups is the number of divisions of the total population. NTotal is the population size. Subsequently, the recovery rate (γ), reproduction number (R_0), incubation rate (δ), percentage of asymptomatic infected people (μ) and detection of asymptomatic (η) values that were previously mentioned in the mathematical model in Section 5.3. Also, the population size (N), initial susceptible individuals (S_0), exposed individ-

uals (E_0), asymptomatic individuals (A_0), infected (I_0) and recovered population per each subgroup from 1 to n separated by one space. Finally, the contact matrix between each subgroup including also the inner contact rate for each one. For example in the input file in figure 7.1 the line after *s1* conform the inner contact as 0.1, contact with subgroup 2 as 0.4 and with subgroup 3 as 0.5 which implies that a 50% of the population in subgroup 1 visits the subgroup 3.

7.1.2 Data Selection

To select instances to evaluate our approach, data from countries with high Coronavirus testing and variety of population and subdivision was selected.

The number of tests and daily cases of countries were obtained from *ourworldindata.org* [14] and open access and open source databases for research and media, and for the Chilean metropolitan region from *coronavirus.mat.uc.cl* [9], COVID-19 data from the Chilean ministry of science and technology and processed for Data UC. The chosen data and consequently the created instances for this work are the following:

- **Denmark:** subdivided into five regions, population of 5,840,045 and 2,841.41 COVID-19 tests per 1,000 people.
- **Qatar:** subdivided into eight municipalities, population of 2,404,776 and 528.21 COVID-19 tests per 1,000 people.
- **Austria:** subdivided into nine states, a population of 8,935,112 and 1,577.72 COVID-19 tests per 1,000 people.
- **Belgium:** subdivided into eleven provinces, a population of 11,431,406 and 809.87 COVID-19 tests per 1,000 people.
- **United Kingdom:** subdivided into twelve regions, population of 63,285,145 and 1,265.72 COVID-19 tests per 1,000 people.
- **Chile:** subdivided into 16 regions, population of 17,574,003 and 481.44 COVID-19 tests per 1,000 people.
- **Italy:** subdivided into 20 regions, a population of 60,359,546 and 648.80 COVID-19 tests per 1,000 people.

Input.txt

```
#Number of Periods
5
#Period Timelapse, days
30
#NVaccines
100
#NumberOfSubgroups
3
#NTotal
900
#Gamma
0.08
#R0
5
#Delta
0.2
#Mu
0.2
#Eta
0.05

#N Subgroups
400 300 200
#S0 Subgroups
395 300 200
#E0 Subgroups
12 0 0
#A0 Subgroups
0 0 0
#I0 Subgroups
0 0 0
#R0 Subgroups
0 0 0

#Contact
#s1
0.1 0.4 0.5
#s2
0.5 0.2 0.3
#s3
0.5 0.45 0.05
```

Figure 7.1: Example of an input text file.

- **Chile Metropolitan Region:** subdivided into 52 districts and population of $7,645,626$.

7.1.3 Setting-up

Each problem instance requires the definition of the scenario analyzed. For each case it is required to consider the number of subgroups, the population, recovery rate (γ), reproduction number (R_0), incubation rate (δ), percentage of asymptomatic infected people (μ), detection of asymptomatic (η) and the contact matrix. Next, it is explained how these components were determined for each problem instance.

Subgroups

The number of subgroups is created based on the subdivisions per Country mentioned in the previous subsection 7.1.2. The only exception is the Chilean region selected for the instance of the Chilean metropolitan region that considers the subdivision of its districts.

Population

For each input file the total population and subgroup populations were set based on the population size and subdivision sizes obtained from different sources of their past census. This, in order to set a population value closer to reality. Total populations are listed in section 7.1.2.

Contact matrix

Contact matrices are created using a two step procedure. First, an initial contact matrix is created based on a weighted random graph.

There are different methods like *Erdös-Rényi* or *Barabási–Albert* methods to create random graphs, for more information about different methods to create random graphs the work of Yong Ga [13], Volchenkov et. al. [28], and Vega et. al [27] can be revised. The method used in this work to create a random graph is similar to the *Bianconi–Barabási* model. In this case, it consists of assign to each subgroup one of three types of nodes based on its *betweenness* centrality (how strong connections it has with another subgroups). The first type of node with the highest *betweenness* centrality, the second type with medium *betweenness* centrality and

the third type with the lowest *betweenness* centrality. The number of nodes of each type is computed based on how many main subdivisions the population has. For example, in Chile there is only one main subdivision, the metropolitan region with almost 41.89% of the population from 16 subdivisions, then, there are other six subdivisions corresponding to medium *betweenness* nodes and the remaining subdivisions as low *betweenness* nodes. From the previous distributions, contact values are generated with random values based on their node *betweenness*. Nodes with high *betweenness* will get high values than those with low *betweenness* keeping the sum of all these contacts per subgroup to 1.

For all the countries the inner interaction of the node is set with a value between 0.4 and 0.6 given that a large fraction of the population in these big subdivisions work or visit the same subdivision, not like for the Chilean metropolitan region because there is much more movement between districts or rather between subdivisions.

An example of a contact matrix is shown in figure 7.2. Here each row indicates the contact rate between a given subgroup and all the other subgroups in the problem. In this case, the first row (describing *subgroup 1*) has a contact of 0.5 with itself, 0.15 with *subgroup 2*, 0.3 with *subgroup 3* and 0.25 with *subgroup 4*.

$$\begin{bmatrix} 0.5 & 0.05 & 0.2 & 0.25 \\ 0.09 & 0.55 & 0.35 & 0.01 \\ 0.1 & 0.2 & 0.4 & 0.3 \\ 0.2 & 0.05 & 0.15 & 0.6 \end{bmatrix}$$

Figure 7.2: Example of a contact matrix for four subgroups.

After the contact matrix is created, the values per each subgroup can be nearly to the distribution of contacts with other subgroups but maybe not with the correct subgroups, in order to get closer to the reality it goes through a geographical adjustment step based on their proximity. For example, using the contact matrix shown in figure 7.2 and the geographical distribution of subdivisions in figure 7.3, the adjustment works as following: the first row changes the values [0.5 0.05 0.2 0.25] to [0.5 0.25 0.2 0.05]. This due to the proximity between subdivisions. The second and third rows do not change because their current con-

tact rows are already set according to their proximity. The last row changes from [0.2 0.05 0.15 0.6] to [0.05 0.2 0.15 0.6] because the second subdivision is the nearest to the first subdivision.

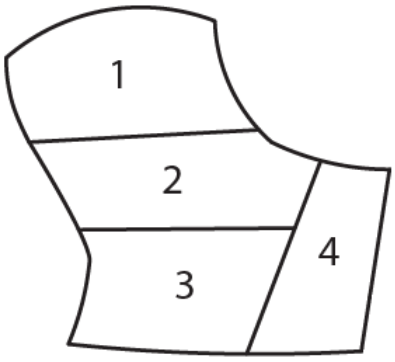


Figure 7.3: Example of country subdivisions for geography adjustment.

$$\begin{bmatrix} 0.5 & 0.05 & 0.2 & 0.25 \\ 0.09 & 0.55 & 0.35 & 0.01 \\ 0.1 & 0.2 & 0.4 & 0.3 \\ 0.2 & 0.05 & 0.15 & 0.6 \end{bmatrix} \rightarrow \begin{bmatrix} 0.5 & 0.25 & 0.2 & 0.05 \\ 0.09 & 0.55 & 0.35 & 0.01 \\ 0.1 & 0.2 & 0.4 & 0.3 \\ 0.05 & 0.2 & 0.15 & 0.6 \end{bmatrix}$$

Figure 7.4: Change of contact matrix in Figure 7.2 after a geography adjustment.

There is a rule for non adjacent subdivisions with the intention of maintaining the *betweeness* of all subgroups. In the case of non adjacent subdivisions if one of them has a higher degree or is a node of high *betweeness* centrality, then the subgroup will get the highest value. For example, if the 4^{th} subdivision has a high preference because has a high *betweeness*, then the resulting matrix from figure 7.2 changes as shown in figure 7.5. In this case, for not adjacent subdivisions as 3 and 4 to 1, the 4^{th} subdivision will get the highest value.

In the case of the metropolitan region the geographical adjustments were made based on mobility data in the metropolitan region in 2012 [15], relationship between the living and work places (data of 2019) [24] and relationship between the living and study places (data of 2018 and 2019) [5].

After these two steps the final contact matrix is obtained for its use in the SEAIRV model.

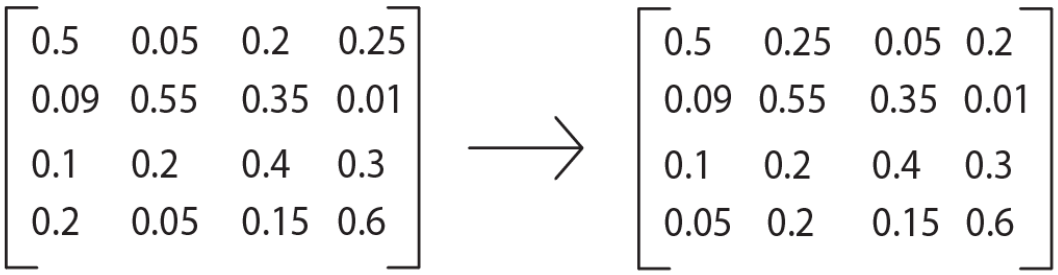


Figure 7.5: Change of contact matrix in figure 7.2 after a geography adjustment considering that 4^{th} region is considered a high degree node.

7.1.4 SEAIR Curve Adjustments per input File

The parameters to adjust to get are divided in two, the COVID-19 based parameters, parameters that depends on the virus. And the SEAIR parameters that were created in this model to resemble as much as possible the reality and are related with the restrictions and movements of the population.

COVID-19 parameters

Once defined the parameters related to the number of subgroups, the population size and the contact rate matrix of section 7.1.3, it is necessary to adjust the parameters of the SEAIR models to the real data curves of COVID-19.

Some of the parameters needed are obtained directly from the disease characteristics. These are the incubation rate (δ), recovery rate(γ) and percentage of infected that are asymptomatic (μ) although they may vary depending on each person. There are other two parameters that are related with the disease, the basic reproduction number R_0, which depends on the way of contact of the disease (air, blood) and the contact of the population, and the detection of asymptomatic (η) that will depend on how Covid-19 testing is performed. For these five parameters, the range values of four of them, δ, γ, μ, R_0, were determined from two sources *Centers of Disease Control and Prevention* (CDC) [22] and *The Centre for Evidence-Based Medicine* (CEBM) [3]. The parameter η depends on each country a range between 5% and 80% was seat.

The ranges of the parameters are shown in table 7.1. The first column correspond to the name of the parameter. The second, fourth and sixth columns cor-

respond to the minimum, average and maximum quantity in days or percentage, each of them followed by the corresponding parameter value.

Parameter	**Minimum**	Value	**Average**	Value	**Maximum**	Value
Delta	2 days	0.5	5-6 days	0.2	14 days	0.07
Gamma	12 days	0.08	2 weeks	0.07	6 weeks	0.024
Mu	10%	0.1	40%	0.5	70%	0.7
Eta	5%	0.05	30%	0.3	80%	0.8
Initial R_0	2	2	3	3	6	6

Table 7.1: Parameter values ranges based on estimated values from CDC and CEBM.

SEAIR parameters

From the SEAIR Model there are six parameters to set for each input file. The parameters are A_1, B_1 and MIP_α for the movement function of the non infected individuals in populations (including asymptomatic) and, A_2, B_2 and MIP_β for the movement function of infected individuals in population. The parameters A_1, B_1, A_2 and B_2 are related with the inclination and variation of the data curve due to small changes in the number of infected. For this, eight possible combinations of the four values that maintain the quantity of infected people is close to the maximum quantity of infected population for the subgroup were seated. The parameters MIP_α and MIP_β define the maximum percentage of population, non infected and infected respectively, allowed to apply a quarantine. Due to this, their values can change between countries according to how restrictive are the measures in terms of population movement.

Adjustments

To set all the previous parameters, it is sought to resemble as much as possible the SEAIR curve with the real COVID-19 data. Figure 7.6 shows a curve of new infected per day. Two curves are shown here. The blue shows the data coming from information systems while the violet shows the SEAIR model curve. The x-axis shows the days and the y-axis the total population. It can be seen how the SEAIR curve and the COVID-19 curve have similarities after the adjustment process. These both epidemic curves have a peak - the highest value - around

17,500 new infected. Both curves show a specific time where a high increase in the number of cases can be identified (close to day 75). Furthermore, the way the curve decreases after the peak can be really prominent reaching values close to zero. In some cases, this decrease can be slower and converge to an equilibrium state of the number of infected people. Curves of different scenarios are different, hence, it is not always possible to get a close-fitting curve to the original one. Despite this, the adjustment process search for finding a set of parameters that lead to the most similar curve.

All the adjusted curves here studied will be presented in chapter 8.

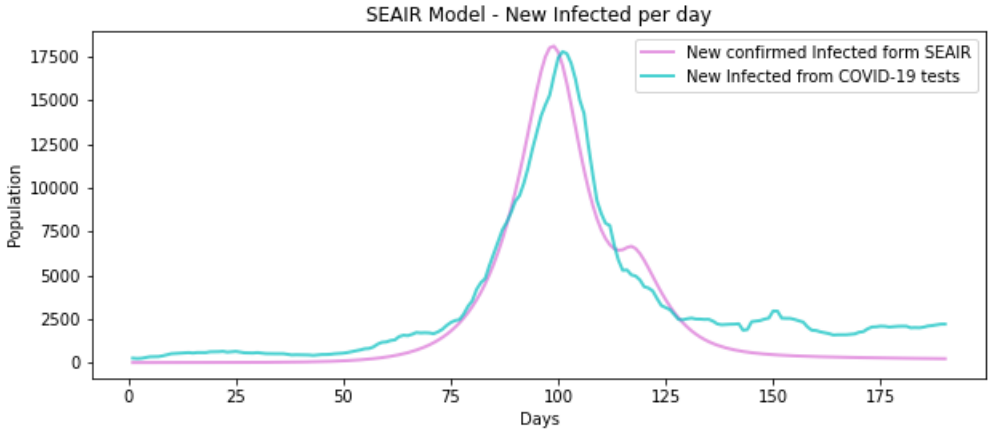


Figure 7.6: Example of new infected persons per day data and SEAIR adjusted curve - Belgium second wave.

It starts seating the COVID-19 parameters in the average values and about the SEAIR parameters the first possible combination of A_1, B_1, A_2 and B_2, and for MIP_α and MIP_β a value of 0.02 and 0.005.

After this, the COVID-19 values are increased or decreased prioritizing the change the R_0 and η values because these are based on features of the population. Also the other three values can be changed but trying to do not set them far away from the average value. These changes are performed until the peak of both curves (Real data and SEAIR model curve) are positioned similarly in the same day (with respect to to the x axis) and also how flat/extended is the curve. Basically, δ, γ and R_0 with highest values move the peak of the curve to the left (because is more exponential in that case) and also made the amplitude higher. About the longitude and amplitude of the curve is related with all the parameters in a greater or lesser way.

With the COVID-19 parameters set before, it is possible to change the MIP_α and MIP_β values, MIP_β determine the maximum quantity of infected or the peak (highest value) of the curve, MIP_α determine how the curve decrease after the peak and also if it get into a equilibrium or tend to zero. The combination of A_1, B_1, A_2 and B_2 parameters give more or less amplitude to the curve, also a big or less fall of the curve after the peak and in case of equilibrium more or less difference with the value of the peak.

The result of these adjustments and the similarity between the SEAIR curve and the COVID-19 tests will be presented for each input file in section 8.

Table 7.2 summarizes all the values that were adjusted to get a curve with the SEAIR model as similar as possible to the COVID-19 data tests. Five of the files are adjusted to the second wave (Denmark, Austria, Belgium, Italy and UK), the other tree (Qatar, Chile and Chilean Metropolitan Region) are adjusted to the first and second wave, this, due to the peak of the curve is higher in the first wave in their case. The first column corresponds to the name of the file, following of the Covid-19 based parameters, recovery rate (γ), incubation rate (δ), reproduction number (R_0), percentage of asymptomatic infected people (μ), detection of asymptomatic (η). After them, the SEAIR function parameters A_1, A_2, B_1, B_2, maximum infected percentage for infected (MIP_β) and non infected (MIP_α) population. Finally, the last two columns of number of vaccines per period and the total percentage of population vaccinated at the end of the five periods.

It is observed that the restrictive values for infected population (MIP_β) are between 0.0100 and 0.0030 and that for non infected population (MIP_α) are between 0.1 and 0.02 of the subdivision population. For the combination of parameters A_1, A_2, B_1 and B_2, the infected curve the values $A_1 = 0.010$ and $B_1 = 1.1$ are the most used with 5 of the 8 input files.For the non infected curve the values $A_2 = 0.001$ and $B_2 = 1.0$ seated for 7 of the 8 input files. Furthermore, it is search the best vaccination plan in the case of covers between 9.48% in the worst case and 31.19% in the best case. The r_0 obtained for the 3 cases that are fitted to their first wave are much higher than the other values, this due to the quantity of tests what were done at the beginning of the pandemic was less, resulting in the need to use a higher R_0 which increase the quantity of infected population much more faster moving the curve to the left.

Input File	γ	δ	R_0	μ	η	A_1	B_1	A_2	B_2	MIP_β	MIP_α	# Vaccines	%
Qatar	0.05	0.15	6.00	0.6	0.20	0.005	1.1	0.001	1.0	0.0055	0.0500	150,000	31.19
Denmark	0.07	0.20	2.25	0.5	0.20	0.010	1.1	0.001	1.0	0.0070	0.0175	250,000	21.4
Austria	0.05	0.12	3.80	0.3	0.50	0.500	5.0	0.001	1.0	0.0064	0.0260	250,000	13.99
ChileMR	0.04	0.10	7.50	0.5	0.40	0.010	1.1	0.001	1.0	0.0065	0.0250	250,000	16.35
Belgium	0.07	0.20	4.50	0.5	0.80	0.010	1.1	0.001	1.0	0.0100	0.0250	300,000	13.12
Chile	0.04	0.10	8.50	0.5	0.20	0.005	1.1	0.01	1.1	0.0030	0.1000	500,000	14.23
Italy	0.06	0.15	5.20	0.5	0.15	0.010	1.1	0.001	1.0	0.0038	0.0200	1,200,000	9.94
UK	0.07	0.20	3.60	0.5	0.30	0.010	1.1	0.001	1.0	0.0052	0.0250	1,200,000	9.48

Table 7.2: Values adjusted for each input file.

7.2 Experiments

In order to set the parameters of the algorithm of the SEAIRV Model explained previously in Chapter 6 three experiments were analyzed. The objective of the first experiment is to evaluate the relevance of the three proposed movements into the local search approach proposed. The second experiment is focused on evaluating the three initialization methods proposed. The last experiment evaluates the two objective functions studied in this approach.

Each experiment was executed five times considering the stochastic nature of the tabu search approach. Each one considering a maximum of 100 iterations and a stopping criteria of 30 iterations without improvement of the best solution.

Four country files with different characteristics were chosen. The inputs are the second wave of Denmark, Austria, Belgium and the first wave of Chile. This countries have 5, 9, 11 and 16 subdivisions respectively. Their maximum quantity of infected per day and shape of their curve are different (The curves are shown in Chapter 8). Also, the second wave of the first tree countries has more tests than the first wave, so their information can be more representative of the country measures. In the case of Chile, the first wave starts almost one month after that in European countries. For this, is it possible that more tests were carried out since the beginning than when the tests were not available yet. Also a scenario where the initial recovered population is zero is an interesting study case.

Because of the time that takes each experiment and the number of tests that should be performed only four of the eight problem instances were chosen.

7.2.1 Movements relevance

For this experiment, we evaluate a set of six probabilities values for the tree movements: Choose a subgroup and give a random quantity of vaccines (*Give Random*), change the quantity of vaccines between two subgroups (*Swap Random*) and invert all the vaccines between two subgroups as a mirror (*Invert Random*) that were explained in Section 6.5.

Each experiment was executed five times. Each one considering a maximum of 100 iterations, a stopping criteria of 30 iterations without improvement of the best solution and initialization of *Inner Interaction*. This process was made for six different parameter configurations. The results are displayed in Section 8.1.

7.2.2 Initialization procedures

The second experiment was for setting which of the initialization process *All to One*, *Inner Interaction*, *Outer Interaction*, *Mixed Interaction* or *Equity* explained in Section 6.3.1 use. Again were made five runs of the algorithm with 100 iterations and a end criteria of 30 iterations without changing the best solution. In this case with the previously set distribution of the movements 80% 10% 10%. The results of this test are in Section 8.1.

7.2.3 Objective functions

The last test is focused into analyzing the factors to optimize, *Maximum Infected* at the same time and *Total Infected* in the period of the objective function explained in Section 6.2. This execution again is to run five times the algorithm with 100 iterations a end criteria of 30 iterations without changing the best solution using the previously seated movement distribution and initialization. The results of it are shown in Section 8.1.

7.2.4 Vaccination plan per Input File

Experiments were made per each input file. As same as the ones before, each experiment was executed five times choosing the best solution obtained from the

five executions. Each one considering a maximum of 100 iterations and a stopping criteria of 30 iterations. Also, each of them with the parameters of movements *Give Random* to 80%, *Swap Random* to 10% and *Invert Random* to 10%. Using the *All to One* initialization and objective functions seated to 50% and 50%. The results of them are shown in Section 8.3.

7.3 Chapter Summary

In this chapter it was explained how the input files were built and what the experiments consist of. Firstly, the format of each input file, after it the precedence of the data and why this data was selected based on the quantity of COVID-19 tests performed, variety of population and subdivisions. After it, how the data of each input file is seated, the subgroups and population from real data and the contact matrix through a two step procedure of the creation of a random contact graph and after this geographical adjustments. Subsequently, how the adjustments are made per each input file, searching for a set of values to COVID-19 parameters and SEAIR parameters. Finally, how the experiments were made. The experiments to set the parameters of the SEAIRV algorithm based on the proposed movements (Give Random, Swap Random and Invert Random), initialization (All to One, Inner Interaction, Outer Interaction and Mixed Interaction) and the objective functions (Minimization of the maximum infected and the total infected).

Results

This chapter presents the results of the experiments carried out. It is composed of the SEAIRV parameter results and the vaccination plans for each input file described in Section 7.1.

First, we present the results obtained for each problem instance using its corresponding SEAIRV parameters. Moreover, we evaluate the relevance of the movements implemented, the initialization procedures and the objective functions for each problem. After this, the results using these parameters are shown by instance, showing the curves obtained with and without vaccination and also the values related to them.

8.1 SEAIRV Parameter Setting Results

These three experiments are executed on only four of the eight problem instances (Denmark, Belgium, Austria and Chile) with a variety of 5, 9, 11 and 16 subgroups were selected. As it was explained in section 7.2 for these experiments only four of the eight files were used due to the time required by experiments.

The results of the three experiments to determine SEAIRV parameters are shown and explained bellow. Each of them has its own results table, description and interpretation.

8.2 Movements relevance

The results from the tests performed to determine the probabilities to perform the three movements: choose a subgroup and allocate a random number of vaccines (*give random*), change the number of vaccines between two subgroups (*swap random*) and invert all the vaccines between two subgroups as a mirror (*invert random*) are shown in table 8.1.

From preliminary experiments we observed that the *give random* movement is the only one that produces changes on the number of vaccines, hence the configurations evaluated always consider a high probability of use.

Input File	Percentage *Give Random*	Percentage *Swap Random*	Percentage *Invert Random*	Average Time [s]	Average Maximum Infected
Belgium	30%	40%	30%	667	166,552.441
	40%	40%	20%	608	166,577.534
	80%	10%	10%	708	166,215.326
	80%	20%	0%	451	166,162.223
	90%	5%	5%	436	166,060.217
	100%	0%	0%	516	**165,720.977**
Austria	30%	40%	30%	303	30,544.340
	40%	40%	20%	329	29,904.129
	80%	10%	10%	311	29,091.669
	80%	20%	0%	373	29,165.058
	90%	5%	5%	297	29,182.102
	100%	0%	0%	337	**28,987.691**
Chile	30%	40%	30%	862	**125,478.953**
	40%	40%	20%	797	125,573.710
	80%	10%	10%	1017	125,847.384
	80%	20%	0%	977	125,760.928
	90%	5%	5%	908	126,026.701
	100%	0%	0%	814	126,185.464
Denmark	30%	40%	30%	92	3,187.872
	40%	40%	20%	109	3,169.145
	80%	10%	10%	100	3,159.761
	80%	20%	0%	94	3,163.694
	90%	5%	5%	101	3,161.494
	100%	0%	0%	119	**3,147.566**

Table 8.1: Average time and maximum infected population obtained per country file for different probabilities of the movements. The best value per file is highlighted in bold.

The first column of the table corresponds to the input file name, the second column is the percentage of the movement *give random*, followed for the percentage of *swap random* and *invert random*. After this, the corresponding average execution time in seconds and average maximum infected as a result of five executions.

Each one considering a maximum of 100 iterations and a stopping criteria of 30 iterations without improvement of the best solution.

For three of the four files selected (Belgium, Austria and Denmark), the use of only one movement *give random* shows the best quality, but for Chile it is completely the opposite. In Chile, the worst number of infected was obtained with this configuration. In terms of time there is not an important difference or any observable pattern related to the movements probability, but it can be noticed that depending on the number of subgroups the time taken to obtain the solutions increase.

Despite the fact that in three of the four cases the use of only one movement shows the best performance, the use of only one movements in the search process can result in a poor performance in other problem instances (like Chile in this case). It could be suggested that the use of only one movement do not allow enough diversification during the search process. This can allow a fast convergence to local optima and premature stagnation of the search. The second best configuration for Austria and Denmark cases were obtained by the same combination of probabilities, 80% *give random*, 10% *swap random* and 10% *invert random*. From this, the values of probabilities per movement is taken as 80% 10% 10%.

8.2.1 Initialization procedures

Here, we evaluate the relevance of the initialization process into the quality and time of results. Five initialization methods were proposed: *All to One*, *Inner Interaction*, *Outer Interaction*,*Mixed Interaction* and *Equity*. As in the previous experiment, each configuration was executed five times. Each one considering a maximum of 100 iterations and a stopping criteria of 30 iterations without improvement of the best solution. Also, in this experiment, the probabilities of movement were seated as 80% *give random*, 10% *swap random* and 10% *invert random*. The results of these experiments are shown in table 8.2.

The first column of the table corresponds to the problem instance name, the second column is the initialization method used. After this, the corresponding average time in seconds and average maximum infected as a result of the five executions.

In this case the best value for three of the four instances, Belgium, Austria and

Input File	Initialization	Average Time [s]	Average Maximum Infected
Belgium	All to One	367	**165,184.050**
	Inner Interaction	420	166,215.326
	Outer Interaction	503	166,002.767
	Mixed Interaction	511	166,132.370
	Equity	468	165,994.828
Austria	All to One	232	**28,657.059**
	Inner Interaction	297	29,182.102
	Outer Interaction	296	29,442.154
	Mixed Interaction	284	28,955.630
	Equity	291	29017.459
Chile	All to One	781	126,664.980
	Inner Interaction	908	126,026.701
	Outer Interaction	969	**125,826.583**
	Mixed Interaction	883	126,299.251
	Equity	1,061	125,858.033
Denmark	All to One	52	**3,146.652**
	Inner Interaction	100	3,159.761
	Outer Interaction	78	3,159.930
	Mixed Interaction	75	3,167.349
	Equity	78	3,160.839

Table 8.2: Average time and maximum infected population obtained per country file for different initialization procedures. The best value per file is highlighted in bold.

Denmark, was obtained using the *all to one* initialization procedure. Moreover, in all the cases, the execution time using the *all to one* procedure is lower than the execution times obtained using any other initialization method. This may be due to a faster convergence and/or stagnation of the search process. Despite the time, again for Chile instance, the worst quality is obtained using the best initialization for the other files. Although the result value of Chile case with *all to one* initialization and considering the same reasons presented before, the initialization will be set to *all to one* for the following experiments.

8.2.2 Objective functions

This experiment aims to analyze the performance of the approach with respect to the two objective functions defined: *minimization of the maximum number of infected* and the minimization of the *total infected* in the period. The experiments were executed using the same parameters values of the previous experiments: probabilities of movement were seated as 80% *give random*, 10% *swap random* and

10% *invert random*; and the initialization was set to *all to one*. Table 8.3 shows the results obtained considering different relevance of objective functions.

Input File	Percentage of MI	Percentage of TI	Average Time [s]	Average Function Value	Average MI	Average TI
Belgium	100%	0%	367	165,184.050	165,184.050	487,775.600
	80%	20%	375	229,702.360	165,184.050	487,775.600
	50%	50%	369	326,479.825	165,184.050	487,775.600
	20%	80%	316	423,257.289	165,184.050	487,775.600
	0%	100%	**359**	487,775.600	165,184.050	487,775.600
Austria	100%	0%	232	28,657.059	28,657.059	48,154.451
	80%	20%	230	32,556.537	28,657.059	48,154.451
	50%	50%	**228**	38,405.755	28,657.059	48,154.451
	20%	80%	231	44,254.97	28,657.059	48,154.451
	0%	100%	253	48,154.451	28,657.059	48,154.451
Chile	100%	0%	781	126,664.980	126,664.980	324,734.161
	80%	20%	699	166,278.816	126,664.980	324,734.161
	50%	50%	754	225,699.570	126,664.980	324,734.161
	20%	80%	**630**	285,120.324	126,664.980	324,734.161
	0%	100%	702	324,733.995	126,665.002	**324,733.995**
Denmark	100%	0%	52	3,146.652	3,146.652	6,304.952
	80%	20%	49	3,778.312	3,146.652	6,304.952
	50%	50%	**47**	4,725.802	3,146.652	6,304.952
	20%	80%	50	5,673.292	3,146.652	6,304.952
	0%	100%	48	6,304.952	3,146.652	6,304.952

Table 8.3: Average time, maximum infected as MI, total infected as TI, and average function value per problem instance for different weighing in the evaluation function. The best time per file is highlighted in bold and the best objective function value (if exists) is shown in bold.

The first column of the table corresponds to the problem instance, the second and third columns show the percentage in the evaluation function to minimize the *maximum number of infected* at the same time and to minimize the *total infected* in the period. After this, the corresponding average time in seconds, average evaluation function value (obtained from the equation in section 6.2), average maximum infected and average Total infected as a result of the five executions.

In this case three of the files do not change their average values on both objective functions, but they get different execution times. This can suggest that the velocity of converge depends on the evaluation function. In the case of Chile, it was possible to get a different solution showing a lower number of total infected and a higher maximum infected at the same time, this demonstrates that cases with higher infected at the same time can have lower quantity of total infected. For two of the files (Austria and Denmark) the best execution times were obtained with percentage 50% and 50%.

For the remaining experiments we used a 50% 50% in order to consider both objectives having the same relevance.

8.3 Vaccination plan per problem instance

In this section, we show for each problem instance its final adjusted curve to their COVID-19 tests data. Furthermore, the best vaccination plan obtained from previous experiments, is applied to the case and the changes between the scenarios with and without vaccination are calculated and the new adjusted curve is displayed.

8.3.1 Qatar

First and Second Wave

In order to apply the SEAIRV model, and adjustment between curves was done. For Qatar, the model was adjusted to fit both, its first and second waves. The curve of new infected per day after these adjustments is presented in figure 8.1 and the cumulative infected population in figure 8.2. For these both graphs the x-axis shows the period of time (days) and the y-axis shows the number of individuals.

In this case it can be observed that the *equilibrium* value of new infected per day from the real data curve is located close to 250, but for the SEAIR curve is closer to 750 new infected. Moreover, in this case it was necessary to consider a *vacation period* in order to relax the movement measures and show a new increase of the infected population. This period was set from day 320 to 350 with an increase of 0.001 to infected and non infected movement rates. From the cumulative infected population, it is observed that both curves begin to flatten after day 80.

From the proposal execution, the vaccination plan conformed of 150,000 vaccines per period is obtained. Part of it is showed in table 8.4, this table shows the six subgroups to which a higher number of vaccines were allocated.

The first column of the table shows the period of vaccination, all other columns show subgroup considered. In this case six of the eight municipalities of Qatar: Ad Dawhah, Al Rayyan, Al Wakrah, Al Khor, Al-Shahaniya and Al Shamal.

It can be observed that in three of the five periods all the vaccines were dis-

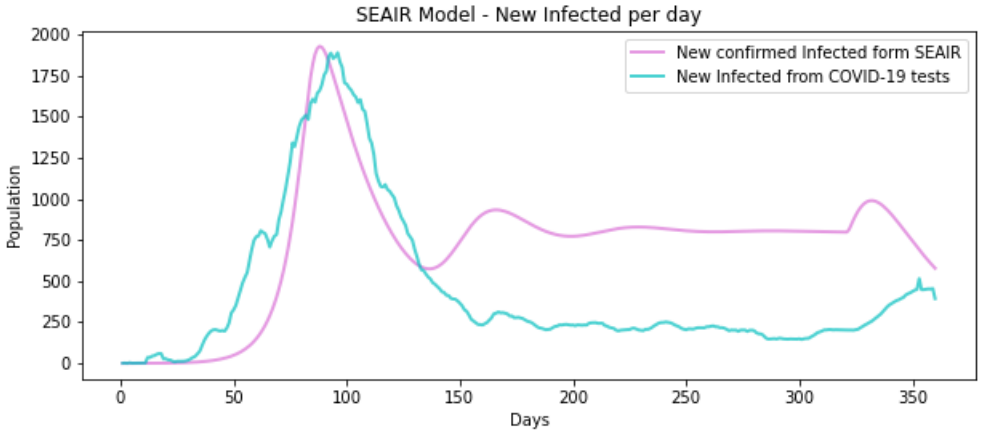


Figure 8.1: Comparison of new infected people between SEAIR Model and COVID-19 tests in Qatar.

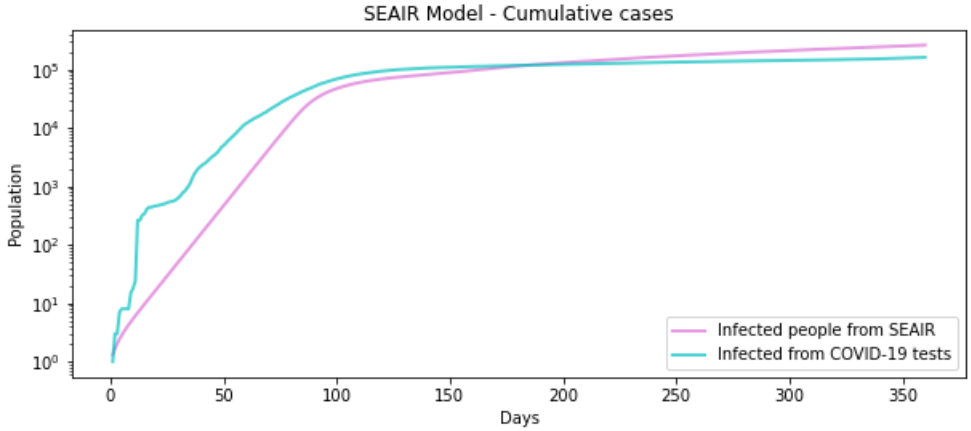


Figure 8.2: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Qatar.

tributed to only one subgroup. Also Al Rayyan gets a higher number of vaccines at the end of the five periods.

	Ad Dawhah	Al Rayyan	Al Wakrah	Al Khor	Al-Shahaniya	Al Shamal
1^{st} Vaccination	0	0	0	150,000	0	0
2^{nd} Vaccination	32,436	102,958	14,606	0	0	0
3^{rd} Vaccination	0	150,000	0	0	0	0
4^{th} Vaccination	39,392	0	2,794	51,496	47,032	5,542
5^{th} Vaccination	0	150,000	0	0	0	0

Table 8.4: Vaccines corresponding to the most vaccinated subdivisions of the best vaccination plan obtained from SEAIRV Model.

The result curve after the vaccination process is shown in figures 8.3 and 8.4.

Here x-axis shows the days and y-axis show the number of individuals.

From the new infected per day curve it can be observed that the peak decreases considerably from more than 1,750 to less than 1,400, also that the *equilibrium* state starts before and, instead of a complete equilibrium, it shows a slow decrease. After the vaccination plan, the vacation period noticeably decreases. In these results the vacation period does not affect the equilibrium state. From the cumulative cases curve, it is noticeable that at the end, they converge to similar values, this is because of the *equilibrium* state that it is still higher than the real value, but it is significantly lower than without vaccination.

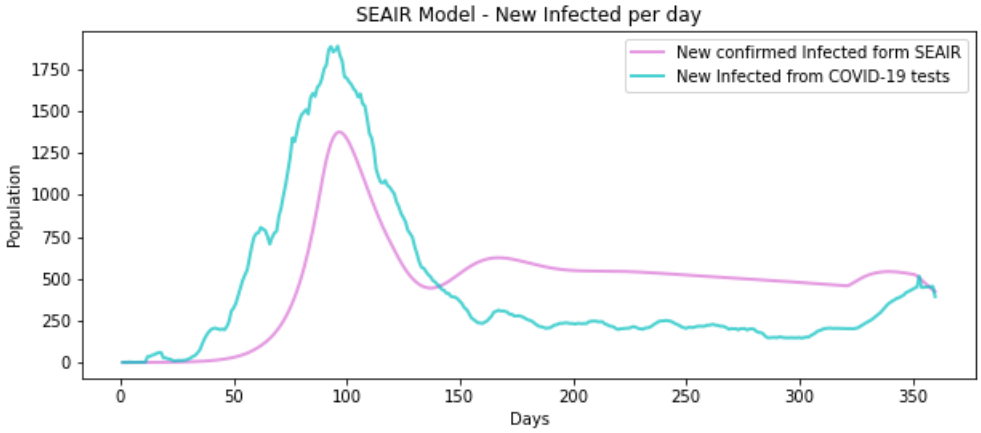


Figure 8.3: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Qatar. First and second wave after applying the best vaccination plan found.

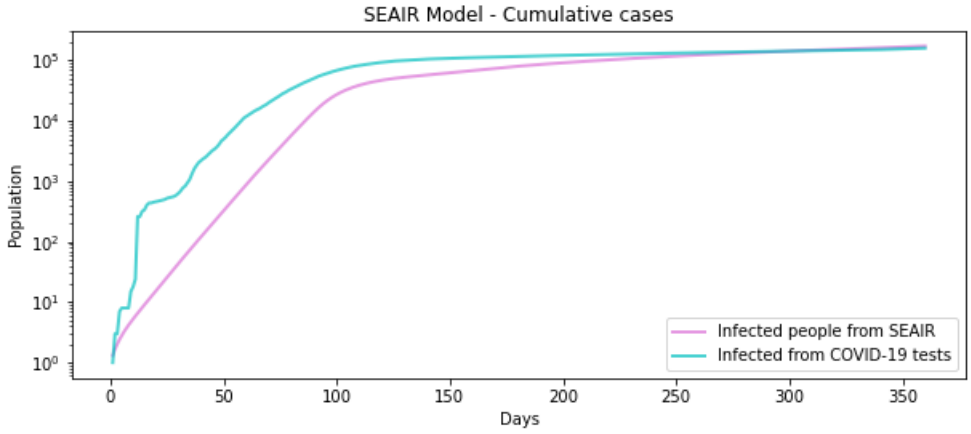


Figure 8.4: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Qatar. First and second wave after applying the best vaccination plan found.

Figure 8.5 shows the curves of each of the subdivisions in Qatar when the best vaccination plan found is applied. In these figures, the number of susceptible (S), exposed (E), asymptomatic (A), infected (I), recovered (R) and new infections over time are presented. In all these plots the x-axis shows the period of time (in days) and the y-axis shows the number of individuals.

It can be observed from the susceptible population that the subgroup with more population get vaccines but not more than the second subgroup with more population. This may be due to how connected is Al Rayyan compared to the largest subgroup Ad Dawhah that after vaccination can generate a lower number of new infections or, that the disease spread slower.

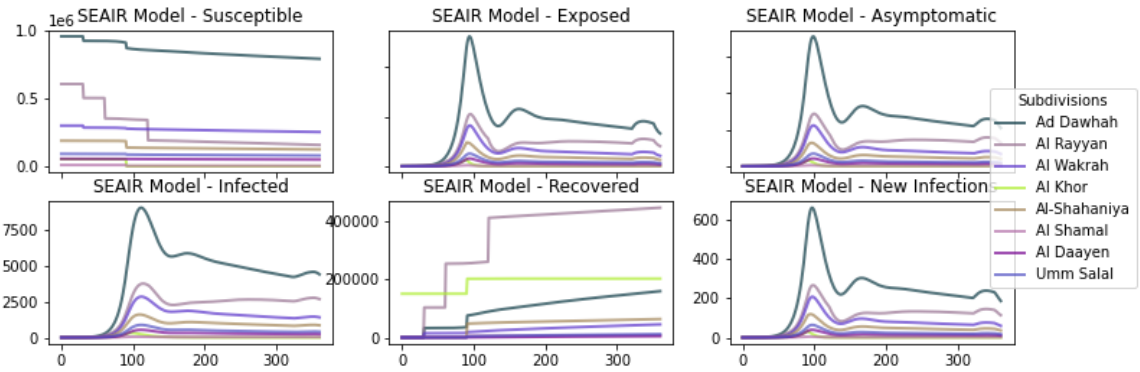


Figure 8.5: Plots of Susceptible, Exposed, Asymptomatic, Infected, Recovered and New infections per day for each subgroup of the best vaccination plan found.

The results of the vaccination plan compared to not vaccination are shown in table 8.5.

The first column of the table corresponds to the evaluated criterion: the total infected population, maximum number of infected and asymptomatic at the same time, maximum number of infected at the same time and maximum new infected per day (or the peak of the curve). Next columns show the corresponding result with the best vaccination plan, without vaccination and the corresponding perceptual gain.

This vaccination plan covers about a 31.19% of the total Qatar population and it is observed that, in this case, the decrease percentages are similar to this value. The total infected is the value that more decreases when vaccination is applied. The maximum number of infected and asymptomatic is the criterion that decreases the less.

	With Best Vaccination	Without Vaccination	Decrease Percentage
Total Infected	173,654.049	260,542.683	33.349%
Maximum Infected and Asymptomatic	21,678.786	29,057.292	25.393%
Maximum Infected	18,808.101	25,243.996	25.495%
Maximum new infected per day	1,375.461	1,926.687	28.61%

Table 8.5: Comparison table of infections with and without vaccination.

8.3.2 Denmark

First Wave

Owing to the data from the first wave in European countries has not enough tests, it is harder to adjust the curve to the data as you can be seen in figure 8.6. Also, in the curve of cumulative cases of this first wave it is noticeable the leak of information at the beginning of the period and how COVID-19 testing started to increase in time.

In figures 8.6 and 8.7 it can be observed the resulting curve of new infected per day and cumulative cases respectively, after the adjustments to the data of COVID-19 test of Denmark. The x-axis shows the days since the start of the disease and the y-axis shows the number of individuals. To get a similar curve for this first wave it was necessary a basic reproduction number (R_0) with a high value of 9.5

From the infected per day curve it can be observed that after the peak it decreases in a gradual way. Also from the cumulative curve it can be seen the leak of information at the beginning of the period with some notorious drops and rises in the cumulative cases.

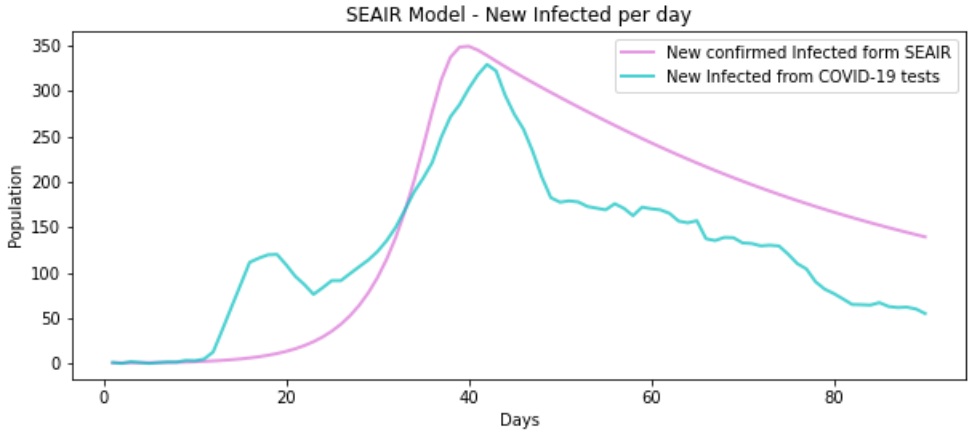


Figure 8.6: Comparison of new infected people between SEAIR Model and COVID-19 tests in Denmark. First wave.

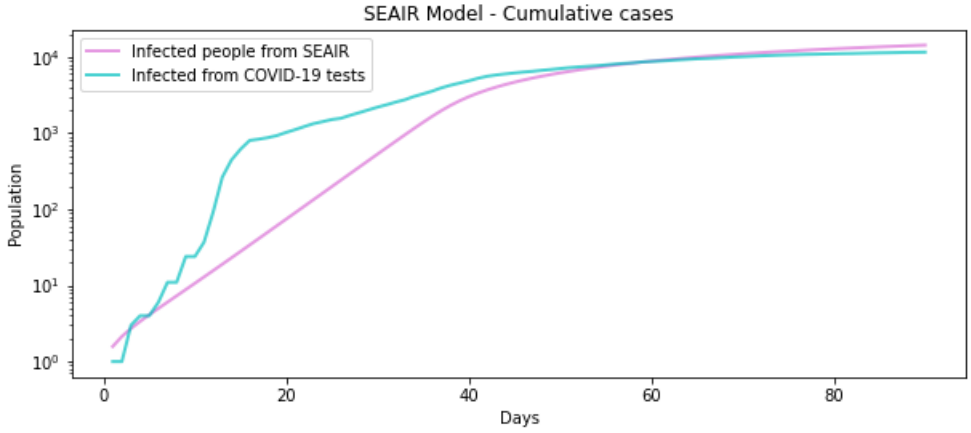


Figure 8.7: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Denmark. First wave.

Second Wave

In order to apply the SEAIRV model an adjustment between curves was made. From this, the figures 8.8 and 8.9 show the curve of new infected per day and cumulative cases of the second wave. The x-axis shows the days since the start of the disease and the y-axis shows the number of individuals.

It can be observed that the peak of the new infected per day curves is almost the same, also after the peak there is a decreasing of both curves to a close value. For the cumulative cases again the curves get really close and show a moderate slope between days 100 and 200.

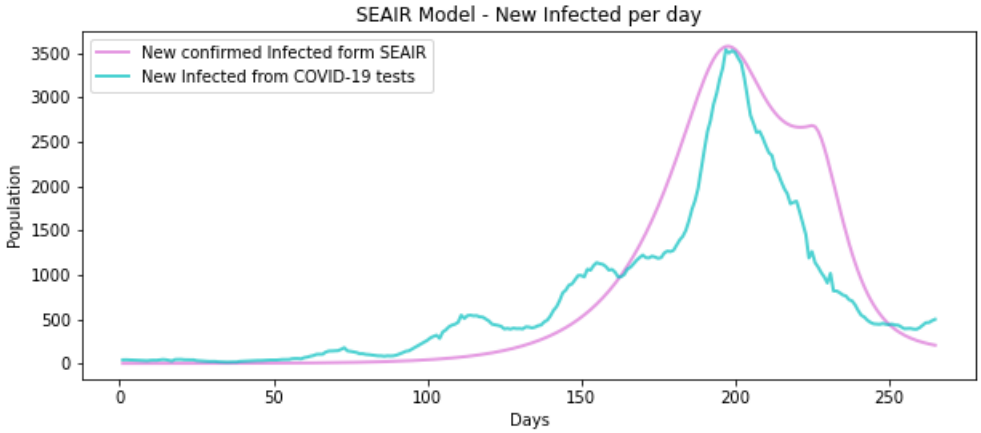


Figure 8.8: Comparison of new infected people between SEAIR Model and COVID-19 tests in Denmark second wave.

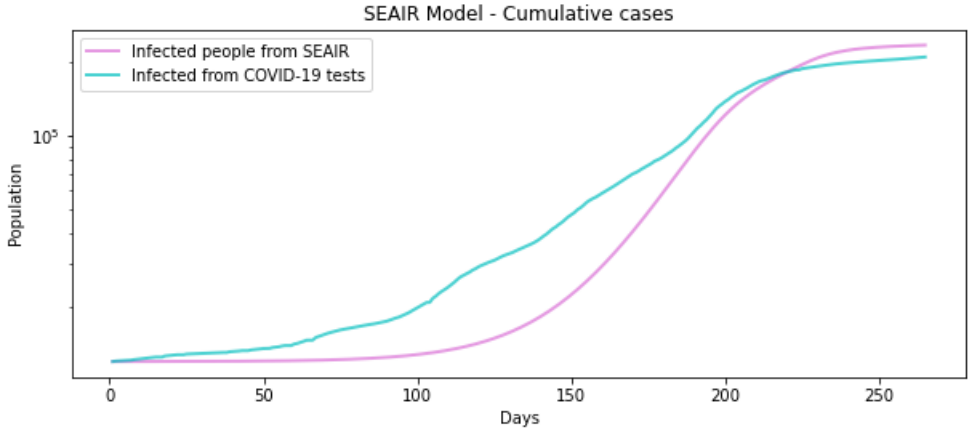


Figure 8.9: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Denmark second wave.

Once verified that the curves are similar, it is possible to proceed to execute the algorithm. In this case to vaccinate 250,000 people each 30 days five times, or vaccinate until the day 150.

The algorithm executes six times with 150 iterations a stopping criteria of 40 iterations with no improvement and the parameter values set in section 7.2.1. The best obtained vaccination plan is shown in table 8.6.

The first column of the table corresponds to the period of vaccination, the next columns correspond to the name of the subgroup. In this case Hovedstaden, Midtjylland, Syddanmark, Sjælland and Nordjylland are the five regions of Denmark.

	Hovedstaden	Midtjylland	Syddanmark	Sjælland	Nordjylland
1^{st} Vaccination	0	0	0	0	250,000
2^{nd} Vaccination	168,060	56,363	419	25,158	0
3^{rd} Vaccination	14,360	0	196,157	39,483	0
4^{th} Vaccination	179,988	15,631	32,347	22,034	0
5^{th} Vaccination	13,176	124,383	41,548	70,893	0

Table 8.6: Best vaccination plan obtained from SEAIRV Model.

The first vaccination allocates all the vaccines to the subgroup that starts with infected population. For this case, the problem instance is the one with lowest basic reproduction number (R_0) implying that the virus spreads slower than in the other cases. At the beginning of the first period, between day 0 and 30, the infected population started in Nordjylland and because of the instance characteristics does not spread enough to another subgroups. For the other vaccination periods, the distribution is more varied, this can be because the infected population starts spreading faster and being influenced by the contact between subgroups. Also, it is observed that Hovedstaden is the one that gets more total vaccines, this probably because it is the subgroup with more population.

In figures 8.10 and 8.11 the resulting curves after the vaccination plan is displayed. Again, the x-axis shows the days since the start of the disease and the y-axis shows the number of individuals. It is observed how the curve of new infected per day is now much more flat and grows slower generating a shift on the position of the peak. Furthermore, in the cumulative cases curve, it is noticed how the maximum quantity of infected cases decrease with vaccination and also how the slope has a lower inclination.

In figure 8.12 it can be observed the curves of each of the subgroups in Denmark after the best vaccination plan found by the SEAIRV algorithm is applied. In the plots the number of susceptible (S), exposed (E), asymptomatic (A), infected (I), recovered (R) and new infections over time are presented with x-axis as days and y-axis as number of individuals.

It can be observed from the susceptible population that the vaccination does not follows a similar pattern in time. Also, for three of the five regions, Hovedstaden, Syddanmark and Sjælland, are granted vaccines for four periods consecutively. Moreover, the second region with less population (Midtjylland) obtains a lower number of recovered people. This may be due to this place is not a region that get too infectious or is not a region that infect other regions.

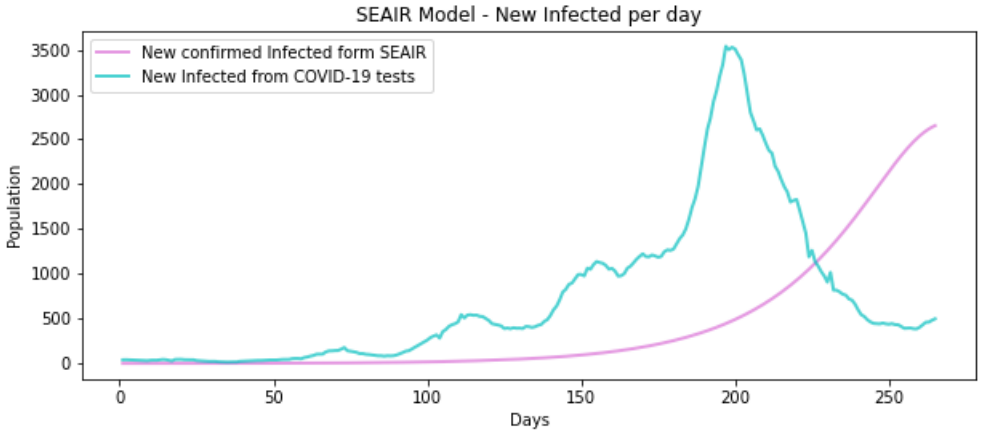


Figure 8.10: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Denmark second wave after applying the best vaccination plan found.

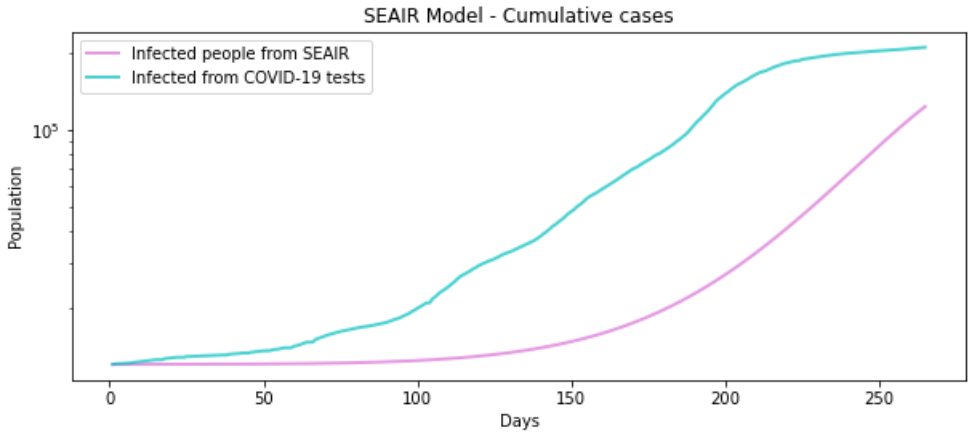


Figure 8.11: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Denmark second wave after applying the best vaccination plan found.

The comparison table between values obtained with and without vaccination are shown in table 8.7. The first column of the table corresponds to the evaluated criterion that can be the total infected population, maximum quantity of infected and asymptomatic at the same time, maximum quantity of infected at the same time and maximum new infected per day (or the peak of the curve). The next columns shows the corresponding results with the best vaccination applied and without vaccination. The last column shows the percentage that the value improves after the vaccination plan is applied.

In this case the total vaccines applied at the end of the periods corresponds to

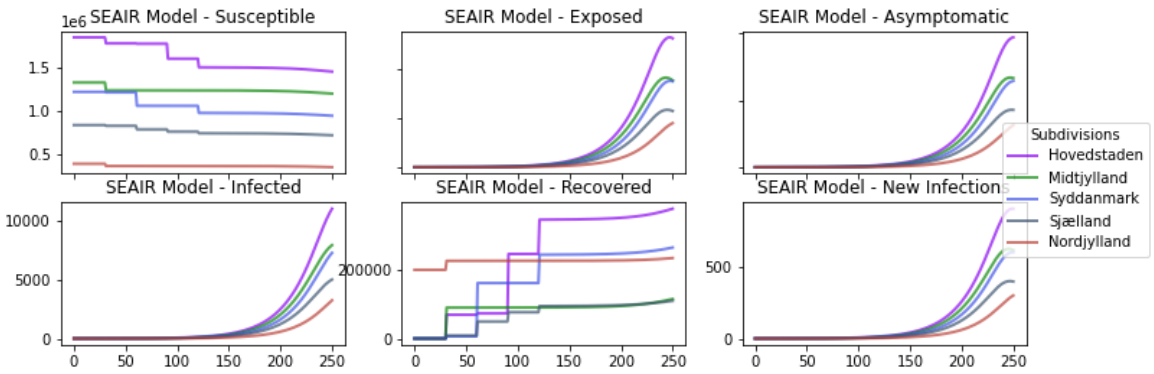


Figure 8.12: Plots of Susceptible, Exposed, Asymptomatic, Infected, Recovered and New infections per day for each Region of the SEAIR Model with the best vaccination plan found.

21.40% of the population. It is obtained a decrease percentage from the total infected of 66.021%. This high value can be due to the fact that after vaccination, the curve gets flatter and reaches its peak value in much more days. The maximum number of infected, asymptomatic and new infected per day decrease in a similar percentage. It is interesting how vaccinating a 21.4% for this instance generates a much important change than the vaccines applied for 31.19% of Qatar population in subsection 8.3.1. This can be owed to the basic reproduction number R_0 higher in Qatar instance and/or the more restrictive measures for non infected population in Qatar.

	With Best Vaccination	Without Vaccination	Decrease Percentage
Total Infected	74,113.695	218,113.191	66.021%
Maximum Infected and Asymptomatic	26,976.820	49,953.341	45.996%
Maximum Infected	22,582.891	43,118.680	47.626%
Maximum new infected per day	2,141.13	3,577.559	40.151%

Table 8.7: Comparison table of infections with and without vaccination.

8.3.3 Austria

First Wave

In order to apply the SEAIRV model, and adjustment between curves was performed. To obtain a similar curve for the first wave with a peak value on approximately day 30, it was necessary a basic reproduction number (R_0) with a really high value of 9.5 this, probably, due to the COVID-19 tests at the beginning of the outbreak were very few.

In figures 8.13 and 8.14 it can be observed the resulting curves of new infected per day and cumulative cases respectively. Once the adjustments of the data of COVID-19 tests of Austria. The x-axis shows the days and y-axis show the number of individuals.

It is observed from the infected per day plot that both curves are similar, their peak is close to 800 and the decay after the peak is a bit slower than the increasing to the peak. From the cumulative cases, both curves are really similar too, getting almost the same value at the end close to 10,000 infected cases.

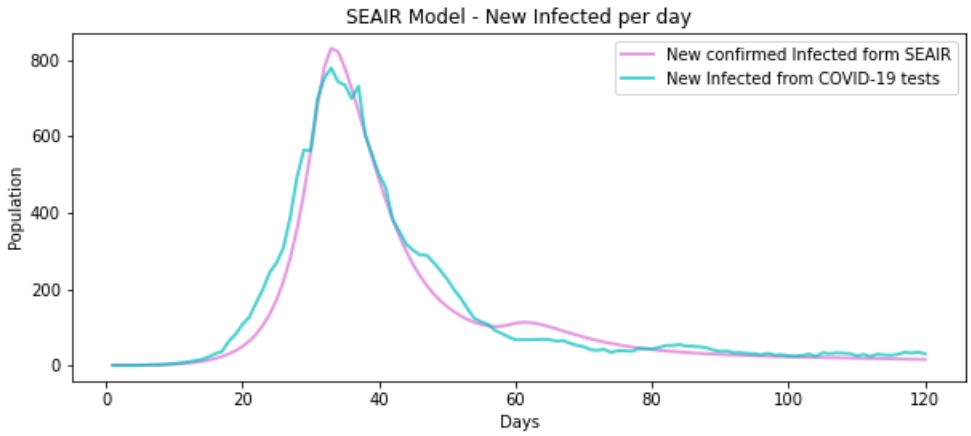


Figure 8.13: Comparison of new infected people between SEAIR Model and COVID-19 tests in Austria. First wave.

Second Wave

Once the adjustment between curves is made for the data of the second wave, the result shown in figures 8.15 and 8.16 of the new infected per day curve and

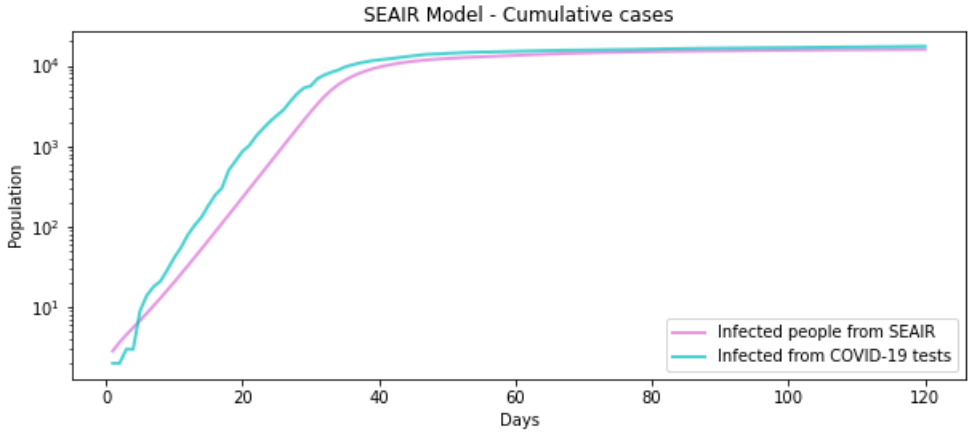


Figure 8.14: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Austria. First wave.

cumulative cases of the second wave. The x-axis shows the days since the start of the disease and the y-axis shows the number of individuals.

It is observed from the new infected cases that, as the opposite of the first curve, the increasing from zero to the peak grows a bit slower and the decay it is more abrupt. Also, from the Cumulative cases near the day of the peak value in both curves they get the same cumulative cases.

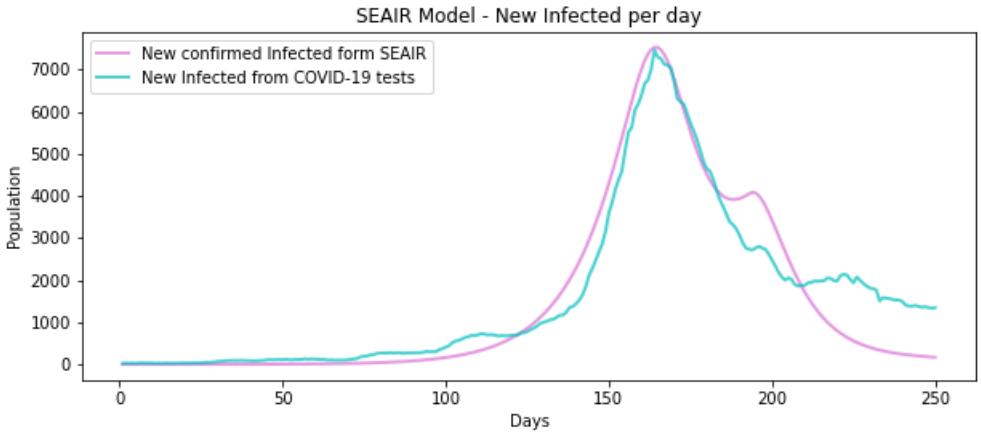


Figure 8.15: Comparison of new infected people between SEAIR Model and COVID-19 tests in Austria. Second wave.

From the proposal execution, the vaccination plan conformed of 250,000 vaccines per period is obtained. Part of it is shown in table 8.8. This table shows the six subgroups to which a higher number of vaccines were allocated.

The first column of the table shows the period of vaccination, all other columns

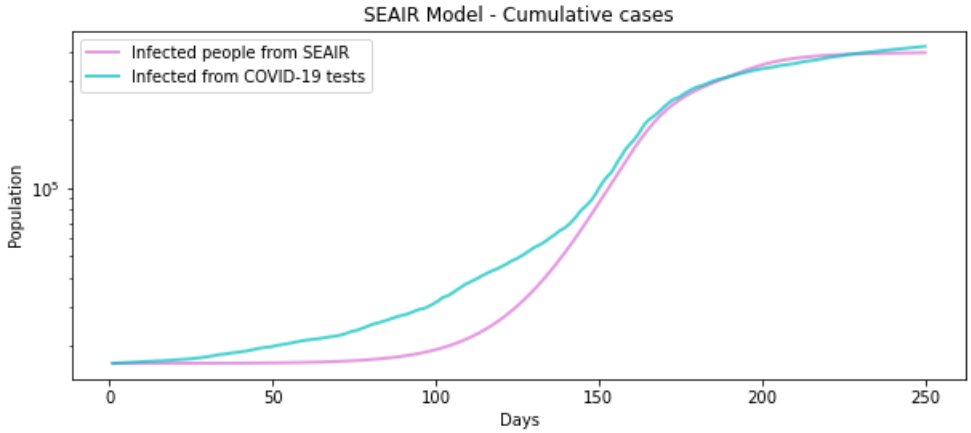


Figure 8.16: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Austria. Second wave.

show the subgroups analyzed. In this case six of the nine states of Austria: Vienna, Lower Austria, Styria, Tyrol, Salzburg and Vorarlberg.

It can be observed that in the first two periods Vorarlberg gets a high number of vaccines. This can be due to, as for Denmark, at the beginning of the outbreak this state had more infected population than others and this number of infected does not increase much until the 3^{rd} vaccination. In the 4^{th} and 5^{th} vaccination Lower Austria and Styria get a high number of vaccines. This can be attributed to how connected these states are to another ones.

	Vienna	Lower Austria	Styria	Tyrol	Salzburg	Vorarlberg
1^{st} Vaccination	0	0	0	0	0	250,000
2^{nd} Vaccination	80,493	4,436	767	3,231	5,138	111,973
3^{rd} Vaccination	3,040	89,984	48,955	8,555	82,470	139
4^{th} Vaccination	0	136,992	23,061	51,267	0	0
5^{th} Vaccination	0	18,918	132,868	71,411	0	0

Table 8.8: Vaccines corresponding to the most vaccinated subgroups of the best vaccination plan obtained from SEAIRV Model.

The resulting curves after the vaccination process is shown in figures 8.17 and 8.18. Here x-axis shows the days and y-axis show the number of individuals. From the new infected per day curve it can be observed that the peak decreases from more than 7,000 to almost 6,000, also compared with the previous SEAIR curve without vaccination of figure 8.15 that decrease to almost zero on day 250. In this case the curve decreases really slow after the peak value. Also, from the cumulative cases, it is observed that the cases start increasing a bit after day 100 compared to the

curve of COVID-19 tests that starts almost at day 0.

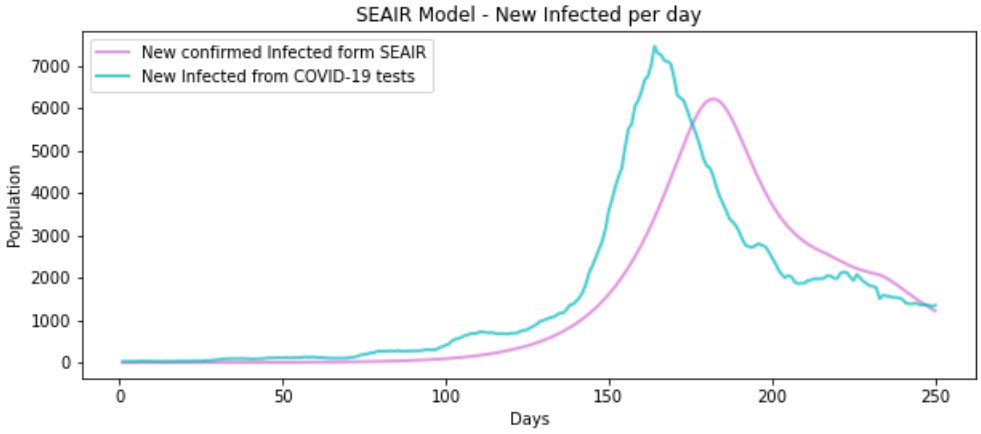


Figure 8.17: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Austria second wave after applying the best vaccination plan found.

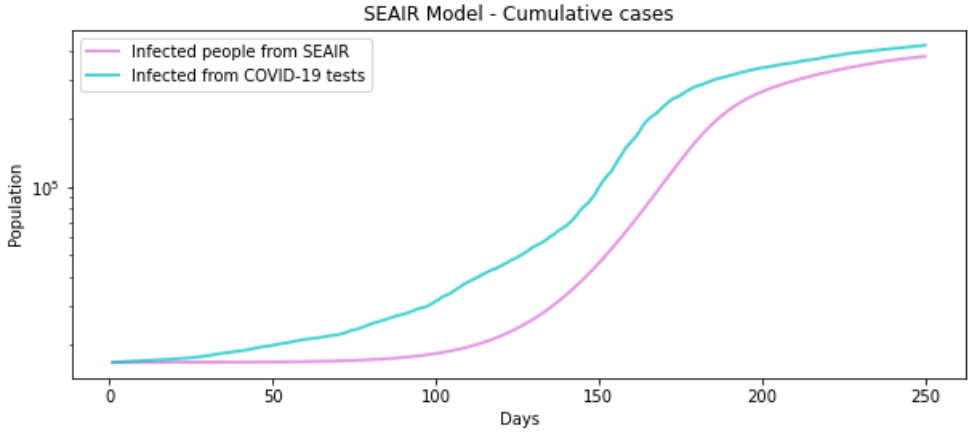


Figure 8.18: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Austria second wave after applying the best vaccination plan found.

Figure 8.19 shows the curves of each of the subgroups in Austria when the best vaccination plan found is applied. In these figures, the number of susceptible (S), exposed (E), asymptomatic (A), infected (I), recovered (R) and new infections over time are presented. In all these plots the x-axis shows the days since the start of the disease and the y-axis shows the number of individuals.

It can be observed from the exposed, asymptomatic, infected and new infected curves how Voralberg subgroup has its peak after day 200. This, when all the other subgroups have their peaks close to day 180. This can be due to the first

and second vaccination applied because the vaccination applied covers more than two-thirds of their population.

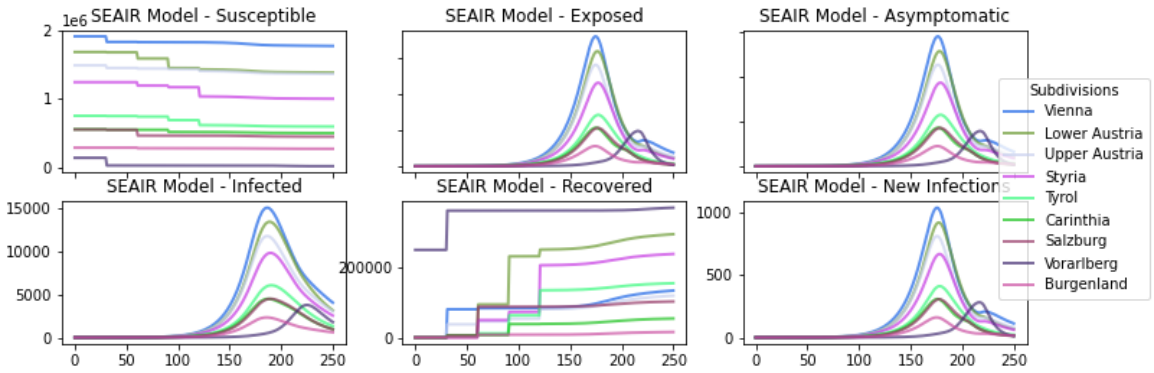


Figure 8.19: Plots of Susceptible, Exposed, Asymptomatic, Infected, Recovered and New infections per day for each Region of the SEAIR Model with the best vaccination plan found.

The results of the vaccination plan compared to not vaccination are shown in table 8.9.

The first column of the table corresponds to the evaluated criterion: the total infected population, maximum number of infected and asymptomatic at the same time, maximum number of infected at the same time and maximum new infected per day (or the peak of the curve). Next columns show the corresponding result with the best vaccination plan, without vaccination and the corresponding perceptual gain.

This vaccination plan covers about a 13.99% of the total Austria population and it is observed that, in this case, the decrease percentages are higher than this value. The total infected is the value that more decreases when vaccination is applied. Regarding to the other criterion they have almost the same decreasing value, this can be owed to the movement restrictions and the quantity of asymptomatic population of 30% being the lower from all the problem instances.

	With Best Vaccination	Without Vaccination	Decrease Percentage
Total Infected	245,494.840	306,179.115	19.820%
Maximum Infected and Asymptomatic	70,114.634	85,351.480	17.852%
Maximum Infected	68,085.025	82,903.757	17.875%
Maximum new infected per day	4,648.093	5,645.949	17.674%

Table 8.9: Comparison table of infections with and without vaccination.

8.3.4 Belgium

First Wave

In order to apply the SEAIRV model, and adjustment between curves was done. To get a similar curve for this first wave with a peak value on approximately day 60 it was necessary to consider a basic reproduction number (R_0) with a value of 5.5 this due to the COVID-19 tests at the beginning of the outbreak were really few.

In figures 8.20 and 8.21 it can be observed the resulting curves of new infected per day and cumulative cases respectively, after the adjustments to the data of COVID-19 test of Belgium were made. The x-axis shows the days since the start of the disease and the y-axis shows the number of individuals.

It is observed from the new infected per day plot that with this data, the curve do not have a pronounced peak, it shows more an equilibrium that last around to 30 days. Also, the curve decay after the peak decreases more softly. From the cumulative cases it can be seen that from day 25 both curves follow almost the same tendency.

Second Wave

In order to apply the SEAIRV model an adjustment between curves was made. From this, the figures 8.22 and 8.23 show the curve of new infected per day and cumulative cases of the second wave. The x-axis shows the days since the start of the disease and the y-axis shows the number of individuals.

It can be observed from the new infected cases that the curve has a very marked

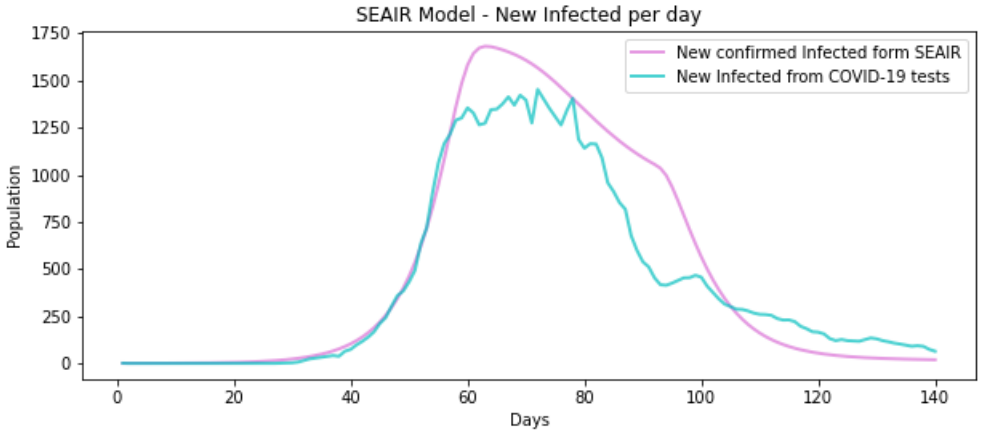


Figure 8.20: Comparison of new infected people between SEAIR Model and COVID-19 tests in Belgium. First wave.

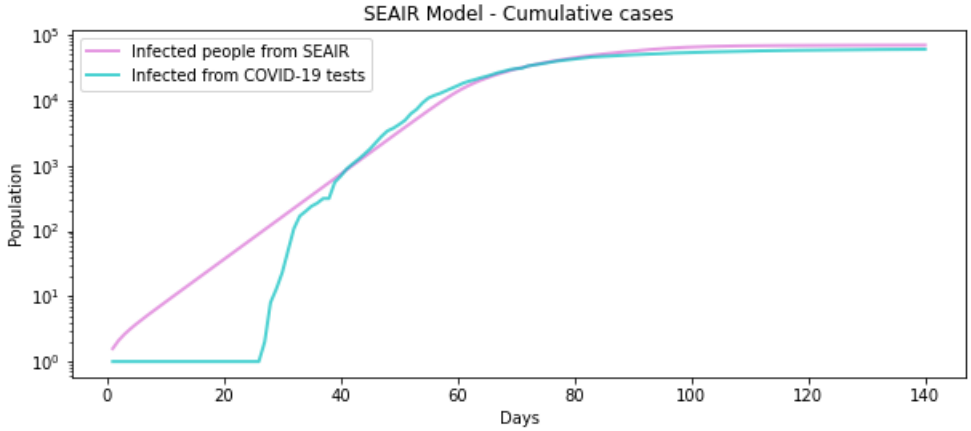


Figure 8.21: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Belgium. First wave.

peak and also a really similar inclination between the increase in cases until the peak and the decreasing of the curve after the peak value. Also that, the COVID-19 tests curve does not end close to zero, rather, it shows an equilibrium state close to 2,500 new infected per day. Regarding to the cumulative cases, both curves have a prominent inclination between days 75 and 100.

From the experiments execution, the vaccination plan conformed of 300,000 vaccines per period was obtained. Part of it is showed in table 8.10, this table shows the six subgroups to which a higher number of vaccines were allocated.

The first column of the table shows the period of vaccination, all other columns show the subgroups considered. In this case six of the eleven provinces of Belgium:

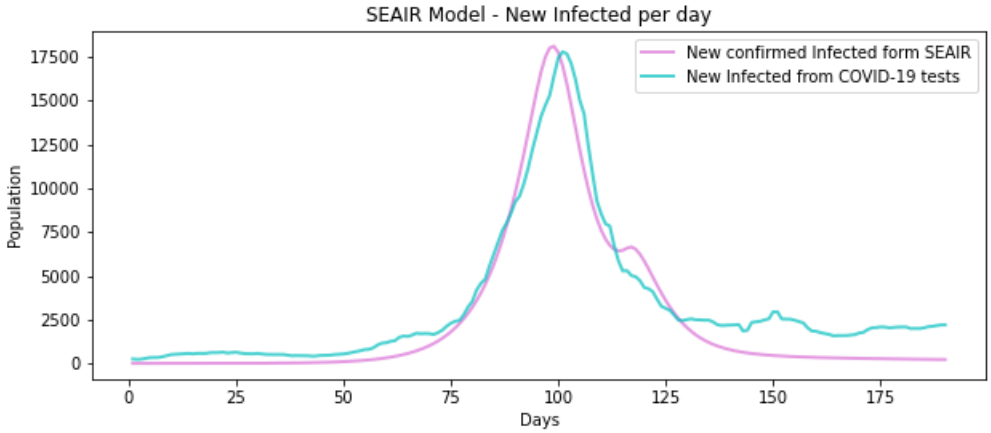


Figure 8.22: Comparison of new infected people between SEAIR Model and COVID-19 tests in Belgium. Second wave.

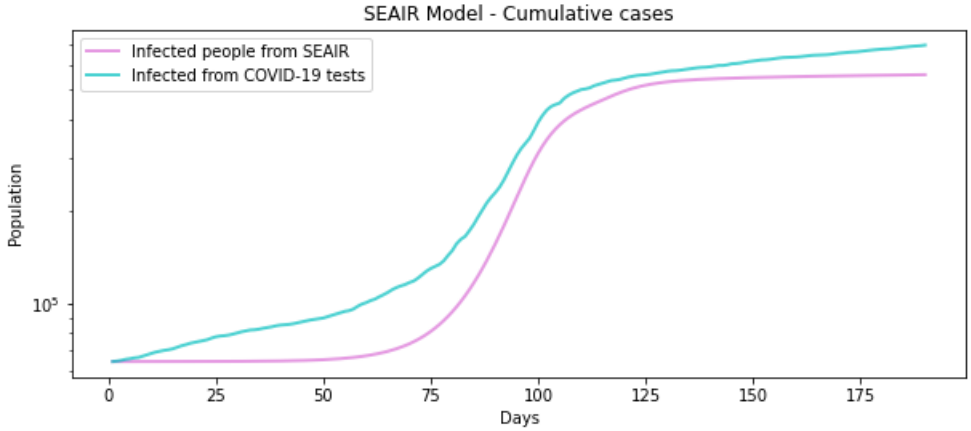


Figure 8.23: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Belgium. Second wave.

Flemish Brabant, Antwerp, Liège, Walloon Brabant, Limburg and Namur.

During the first and fourth vaccination periods all the vaccines were allocated to Antwerp province, but in the other periods there are almost not allocations. A first possible reason can be related to the number of infected population after the vaccination starts decreasing and other subgroups start having a higher amount of infected population. Another possible reason can be attributable to the restriction measures. Antwerp province had too many infected population and individuals here start decreasing their movement reaching the point that a vaccination will not change much the spread of the virus.

In figures 8.24 and 8.25 the resulting curves after the vaccination plan are dis-

	Flemish Brabant	Antwerp	Liège	Walloon Brabant	Limburg	Namur
1^{st} Vaccination	0	0	300,000	0	0	0
2^{nd} Vaccination	91,507	47,449	0	20,133	4,935	10,067
3^{rd} Vaccination	59,017	29,817	16	13,188	83,452	30,376
4^{th} Vaccination	0	0	0	300,000	0	0
5^{th} Vaccination	1,380	2,656	0	62,231	0	22,5805

Table 8.10: Vaccines corresponding to the most vaccinated subdivision of the best vaccination plan obtained from SEAIRV Model.

played. Again, the x-axis shows the days since the start of the disease and the y-axis shows the number of individuals.

It is observed how the curve of new infected per day makes a shift to the right, also the peak decreases its value and the decay after the peak is softer and takes longer to get to zero new infected per day. This, compared to the curve without vaccination obtained with the SEAIR model. The cumulative cases maintains a similar shape, but the values of the SEAIR model are always lower.

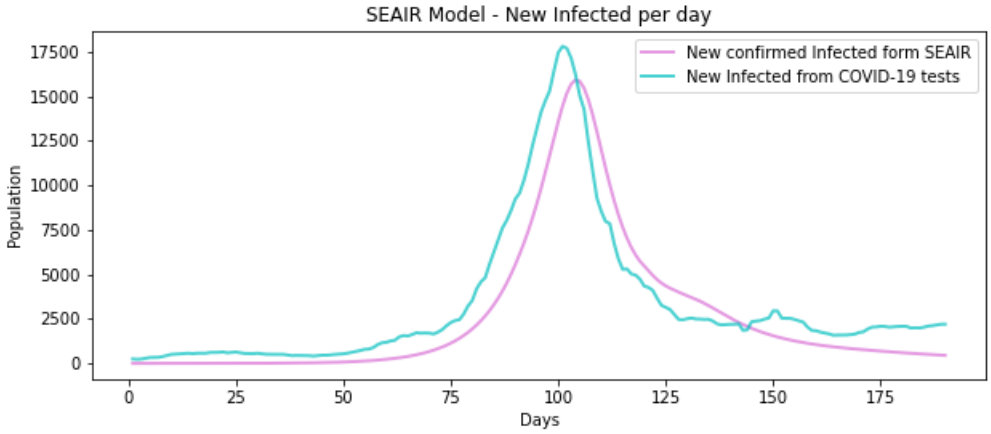


Figure 8.24: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Belgium second wave after applying the best vaccination plan found.

Figure 8.26 shows the curves of each of the subdivisions in Belgium when the best vaccination plan found is applied. In these figures, the number of susceptible (S), exposed (E), asymptomatic (A), infected (I), recovered (R) and new infections overtime are presented. In all these plots the x-axis shows the period of time (in days)and the y-axis shows the number of individuals.

It can be observed that for the exposed, asymptomatic and new infected curves for most of the subgroups have a similar shape with respect to the plot of the first

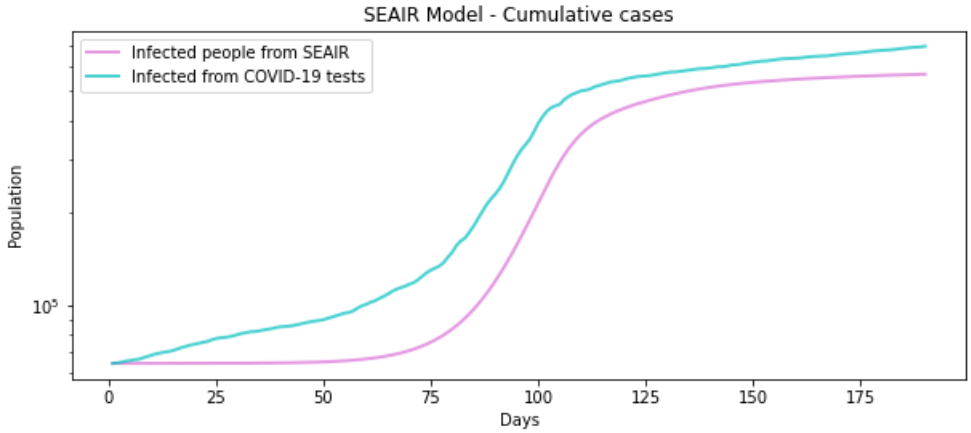


Figure 8.25: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Belgium second wave after applying the best vaccination plan found.

curve without vaccination. But, there are two exceptions the first one corresponds to Wallon Brabant from which you can notice two waves or two peak values due to vaccination. The other exception is Hainaut which has a slower smoother fall from the curve after the peak value that looks more like the infected curve. From the susceptible and recovered plots, it is observed that vaccinations were varied, distributing in each period different number of vaccines to different groups.

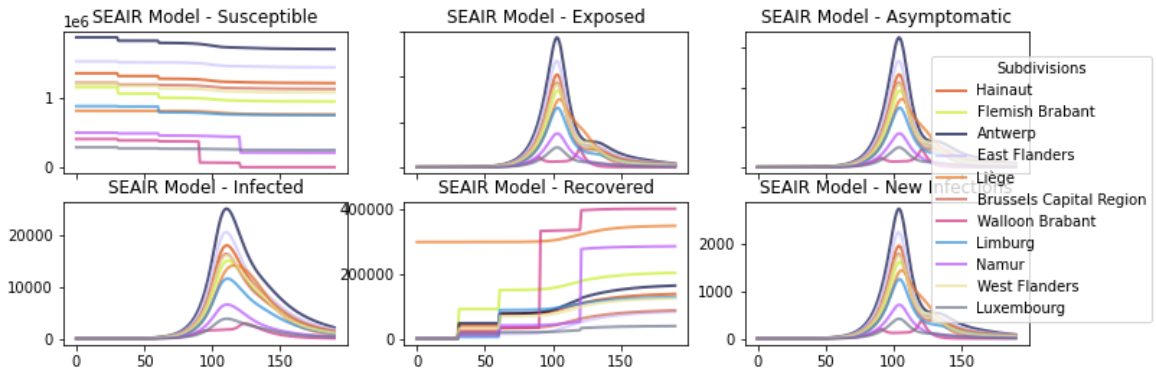


Figure 8.26: Plots of Susceptible, Exposed, Asymptomatic, Infected, Recovered and New infections per day for each Region of the SEAIR Model with the best vaccination plan found.

This vaccination plan covers about a 13.12% of the total Belgium population and it is observed that, in this case, the decrease percentages are lower than this value (table 8.11). Unlike the other vaccination plans a slight increase of the total infected is obtained here. This can be due to the movement factors of the infected population. This instance has the highest movement factor for infected popula-

tion. This can induce a higher value of infected population to completely restrict their movement. By decreasing the maximum new infected per day, more freedom to move to infected population can be allowed.

	With Best Vaccination	Without Vaccination	Decrease Percentage
Total Infected	502,679.096	495,482.655	−1.452%
Maximum Infected and Asymptomatic	155,287.904	169,063.317	8.148%
Maximum Infected	148,644.350	161,669.909	8.057%
Maximum new infected per day	15,914.796	18,116.113	12.151%

Table 8.11: Comparison table of infections with and without vaccination.

8.3.5 Chile

First and Second Wave

In order to apply the SEAIRV model, and adjustment between curves was done. For Chile, the model was adjusted to fit both, its first and second waves. The curve of new infected per day after these adjustments is presented in figure 8.27 and the cumulative infected population in figure 8.28. For these both graphs the x-axis shows the period of time (days) and the y-axis shows the number of individuals.

For this case, a vacation period was applied from day 310 to 370 with a increase in the movement of infected and non infected population of 0.001. From the new infected curve it can be notice that the peak occurs in about day 110, also that the COVID-19 data curve gets a equilibrium state from day 150 to 300 of about 2,000 people per day. The SEAIR model curve reaches its state of equilibrium in about 3,000 people per day. From the cumulative cases after day 100 from both curves values are really similar until day 350.

The solution obtained of he vaccination plan, conformed of 500,000 vaccines per period was obtained. Part of it is shown in table 8.12. This table shows the six subgroups to which a higher number of vaccines were allocated.

The first column of the table shows the period of vaccination, all other columns show the subgroups considered. In this case six of the eleven regions of Chile: Metropolitan, Bío-bío, Los Lagos, Ñuble, Los Ríos and Arica y Parinacota.

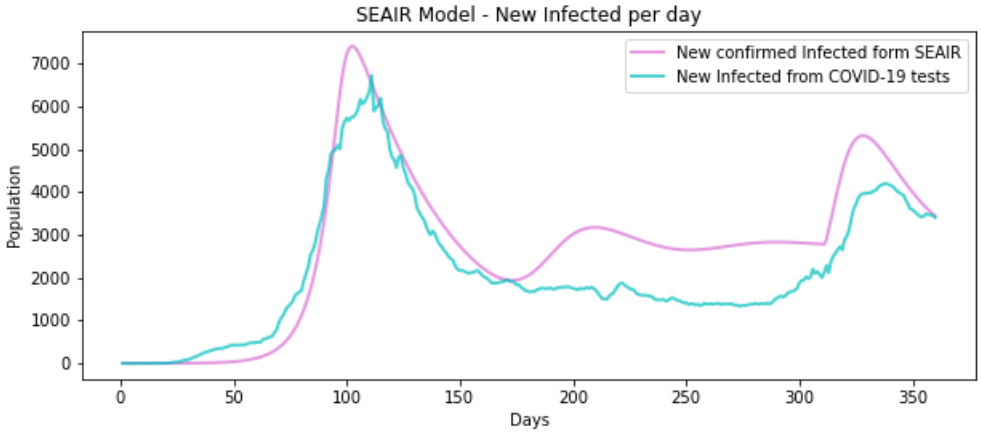


Figure 8.27: Comparison of new infected people between SEAIR Model and COVID-19 tests in Chile. First and second wave.

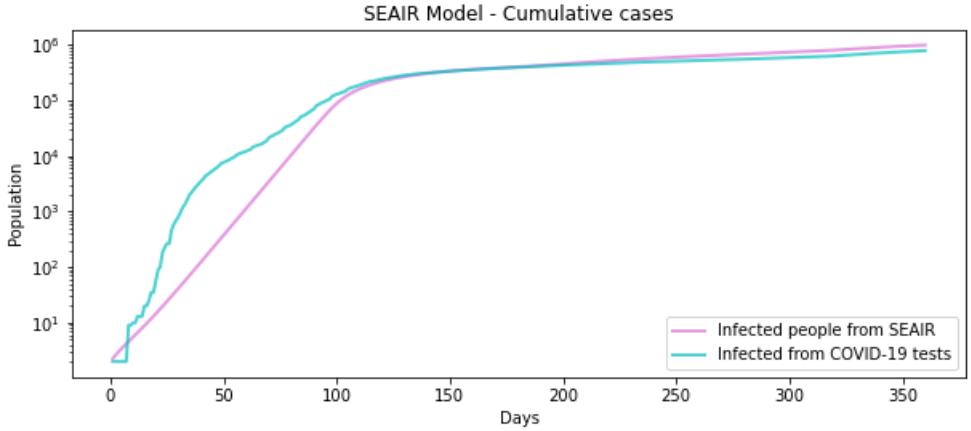


Figure 8.28: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Chile. First and second wave.

It can be noticed that the Metropolitan region and Los Lagos are the subgroups that have more vaccines allocated. The metropolitan region is the subgroup with more population and the most connected. Los Lagos region is the one were the infectious cases started and through of the virus infectiousness, that at the beginning is slow, the first 30 days it maintains at the subgroup with more infected population.

The resulting curve after the vaccination process is shown in figures 8.29 and 8.30. Here x-axis shows the days and y-axis show the number of individuals.

It is observed from the new infected per day that the peak value slightly decrease and the curve shift slightly to the right side. From the cumulative cases a modest

	Metropolitan	Biobío	Los Lagos	Ñuble	Los Ríos	Arica y Parinacota
1^{st} Vaccination	0	0	500,000	0	0	0
2^{nd} Vaccination	292,324	65,504	537	186	2,084	113
3^{rd} Vaccination	410,403	0	0	0	0	0
4^{th} Vaccination	0	52,199	33,2241	94,743	0	0
5^{th} Vaccination	0	0	0	32,033	374,274	93,397

Table 8.12: Vaccines corresponding to the most vaccinated subdivision of the best vaccination plan obtained from SEAIRV Model.

decrease can be observed.

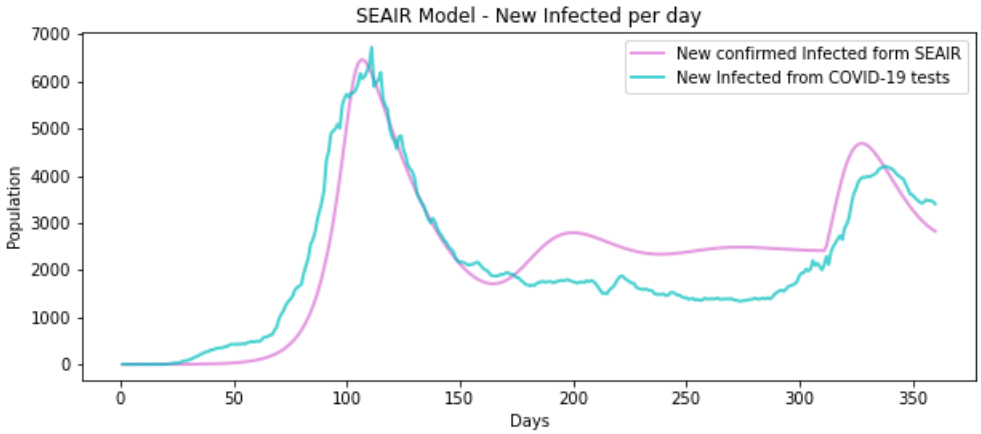


Figure 8.29: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Chile second wave after applying the best vaccination plan found.

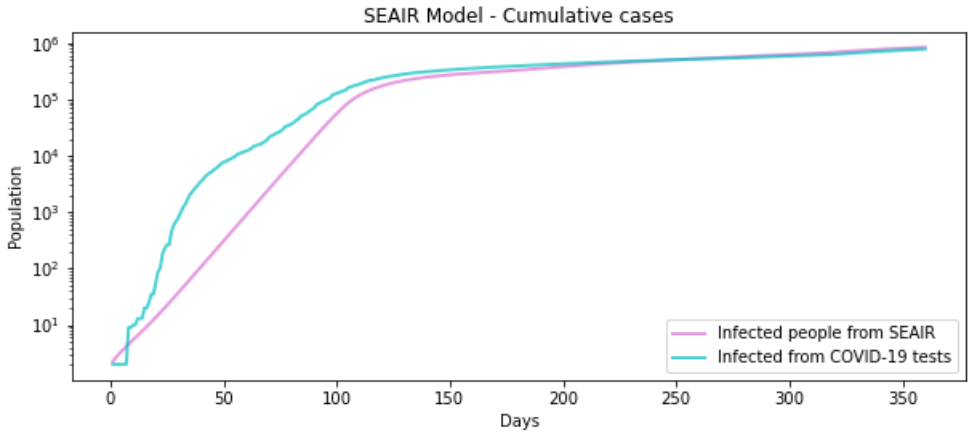


Figure 8.30: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Chile second wave after applying the best vaccination plan found.

Figure 8.31 shows the curves of each of the subdivisions in Chile when the best

vaccination plan found was applied. In these figures, the number of susceptible (S), exposed (E), asymptomatic (A), infected (I), recovered (R) and new infections overtime are presented. In all these plots the x-axis shows the period of time (in days) and the y-axis shows the number of individuals.

It can be notice the big difference between population of the Metropolitan region and other subgroups. Furthermore, it can be seen the difference between the shape of the asymptomatic and infected curves were differences between the peak and the state of equilibrium is much more pronounced in the asymptomatic curve. This can be due to the freedom of movement that asymptomatic population have.

Figure 8.31: Susceptible, Exposed, Asymptomatic, Infected, Recovered and New infections per day for each Region of the SEAIR Model with the best vaccination plan found.

The results of the vaccination plan compared to not vaccination are shown in table 8.13. The first column of the table corresponds to the evaluated criterion: the total infected population, maximum number of infected and asymptomatic at the same time, maximum number of infected at the same time and maximum new infected per day (or the peak of the curve). Next columns show the corresponding results with the best vaccination plan, without vaccination and the corresponding perceptual gain.

This vaccination plan covers about a 14.23% of the total Chile population and it is observed that, in this case, the decrease percentages are similar to this value. Unlike the previous problem instances, the difference between the decrease percentage of total infected and maximum infected is not very important. It should be emphasized that, for this problem instance, different parameter configurations shown a better performance, so it is highly possible that there were much better vaccination plans that deliver a better value for these types of instances.

	With Best Vaccination	Without Vaccination	Decrease Percentage
Total Infected	840,096.735	983,040.848	14.541%
Maximum Infected and Asymptomatic	111,306.477	129,594.850	14.112%
Maximum Infected	101,152.092	117,983.213	14.266%
Maximum new infected per day	6,465.033	7,410.361	12.757%

Table 8.13: Comparison table of infections with and without vaccination.

8.3.6 Italy

First Wave

In figures 8.32 and 8.33 it can be observed the resulting curves of new infected per day and cumulative cases respectively, after the adjustments to the data of COVID-19 tests of Italy. The x-axis shows the days since the start of the disease and the y-axis shows the number of individuals. To get a similar curve for this first wave it was necessary a basic reproduction number (R_0) with a high value of 11. It should be notices that, in this first curve, the COVID-19 tests made were not many and this is the main reason to use a really high value.

From the new infected per day curve it is observed that both curves have similar peak values and also a similar moderate decay after the peak of the curve. After the day 20 of the cumulative cases curve, the values on infected population from both curves present similarities.

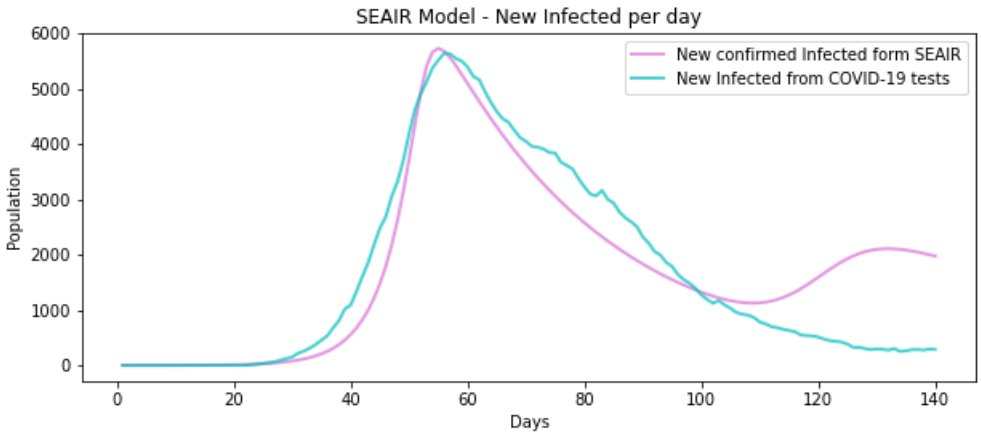


Figure 8.32: Comparison of new infected people between SEAIR Model and COVID-19 tests in Italy. First wave.

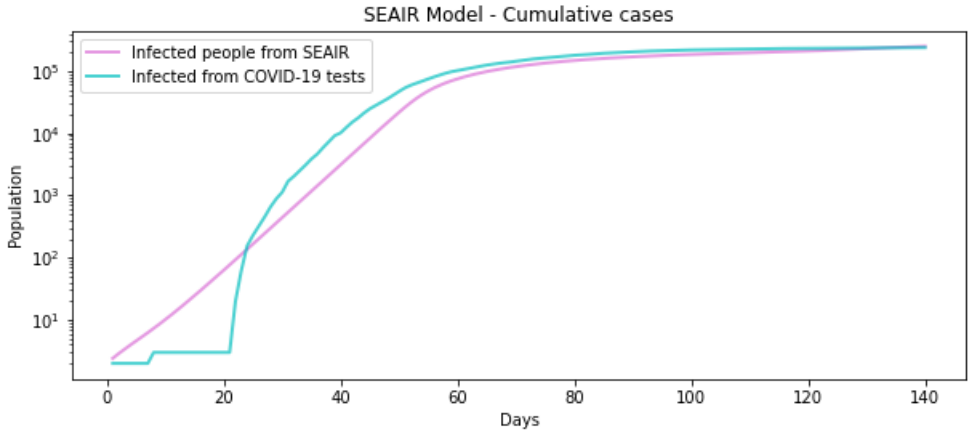


Figure 8.33: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Italy. First wave.

Second Wave

In order to apply the SEAIRV model, an adjustment between curves was made. From this, the figures 8.34 and 8.35 show the curves of new infected per day and cumulative cases of the second wave. The x-axis shows the days since the start of the disease and the y-axis shows the number of individuals.

It can be observed that the peak of the new infected per day for the COVID-19 test curve is locate close to day 125 while for the SEAIR curve it is locate close to day 120. Despite of this, the shape of the curve and the decay before the equilibrium state is similar. From the cumulative cases, both curves have the same shape and similar values at the end of the period in day 175.

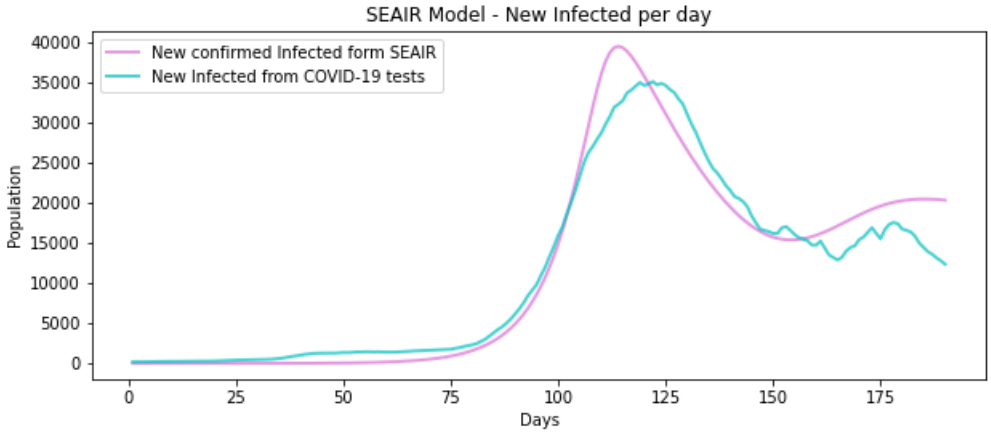


Figure 8.34: Comparison of new infected people between SEAIR Model and COVID-19 tests in Italy. Second wave.

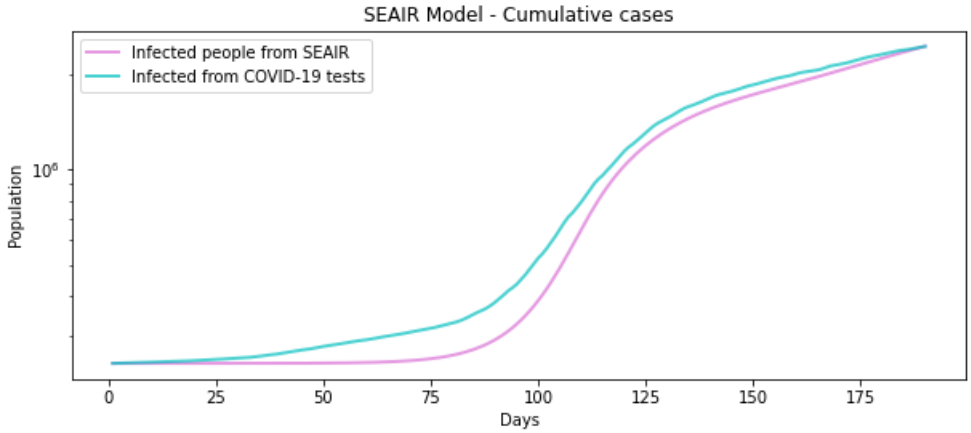


Figure 8.35: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Italy. Second wave.

Once verified that the curves are similar, it is possible to proceed to execute the algorithm. The vaccination plan considered $1,200,000$ vaccines per period. Part of the allocation in shown in table 8.14. This table shows the six subgroups to which a higher number of vaccines were allocated.

The first column of the table shows the period of vaccination, all other columns show subgroups considered. In this case six of the twenty regions of Italy: Piedmont, Puglia, Calabria, Sardinia, Abruzzo and Friuli-Venezia Giulia.

It is observed that in the same way as in other previous instances the first vaccination allocates all the vaccines to one subgroup. Furthermore, it is surprising that Lombardy, the subgroup with more population, is not considered in the set of the most vaccinated. These results may be due to the significance of the first infected when the vaccination started.

	Piedmont	Puglia	Calabria	Sardinia	Abruzzo	Friuli-Venezia Giulia
1^{st} Vaccination	0	0	0	1,200,000	0	0
2^{nd} Vaccination	1,320	0	33,167	496,079	0	669,261
3^{rd} Vaccination	159	3013	610,090	0	195,865	387,350
4^{th} Vaccination	0	0	1,088,912	0	8,600	102,488
5^{th} Vaccination	324,447	212,566	217,312	2,680	0	49,956

Table 8.14: Vaccines corresponding to the most vaccinated subdivision of the best vaccination plan obtained from SEAIRV Model.

In figures 8.36 and 8.37 the resulting curves after the vaccination plan are displayed. Again, the x-axis shows the days since the start of the disease and they-axis shows the number of individuals.

It is observed how the curves of new infected per day decrease the peak value and also it is adjusted a bit to the right side. Furthermore, in the cumulative cases curve, it is noticed how the maximum quantity of infected cases decrease and also the distance between days after the day 75.

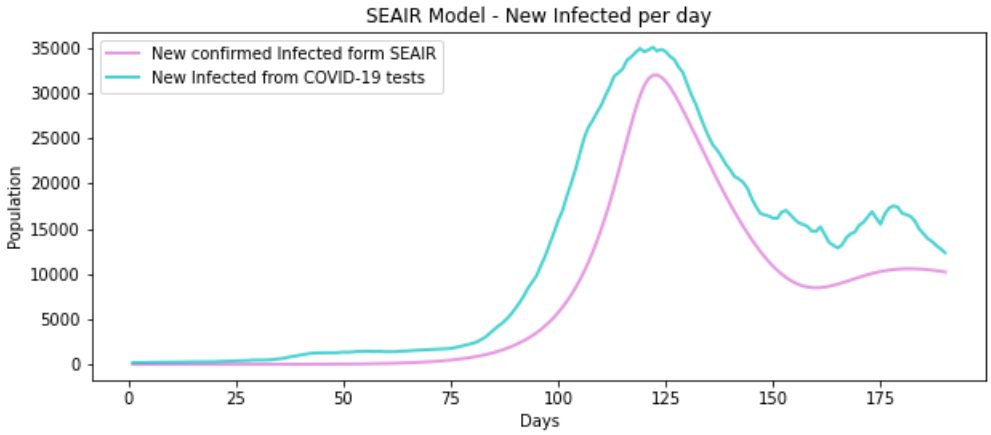


Figure 8.36: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Italy second wave after applying the best vaccination plan found.

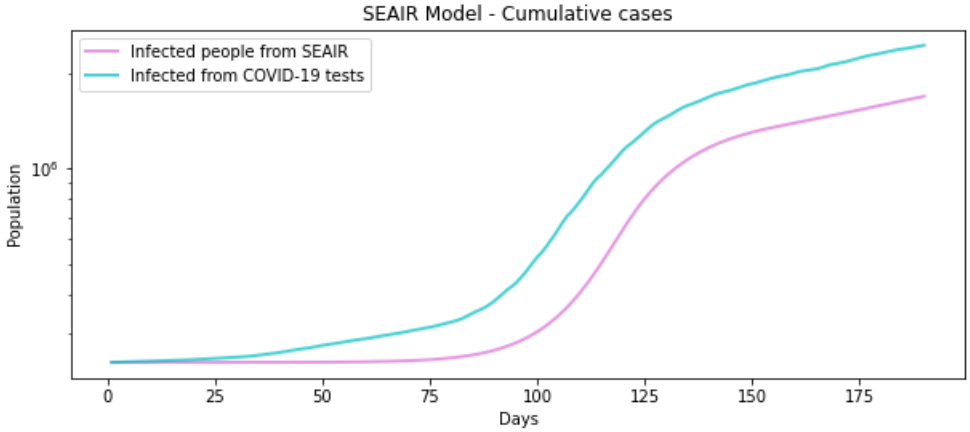


Figure 8.37: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Italy second wave after applying the best vaccination plan found.

In figure 8.38 it can be observed the curves of each of the subgroups in Italy after the best vaccination plan found by the SEAIRV algorithm was applied. In the plots the number of susceptible (S), exposed (E), asymptomatic (A), infected (I),recovered (R) and new infections over time are presented with x-axis as days and y-axis as number of individuals.

It can be notice from the susceptible plot, that most of the vaccinations were made to subgroups with lower population. Also that, as in much of the cases seen before, the asymptomatic, exposed and new infections curve has a greater difference between the peak value and the equilibrium than the infected curve.

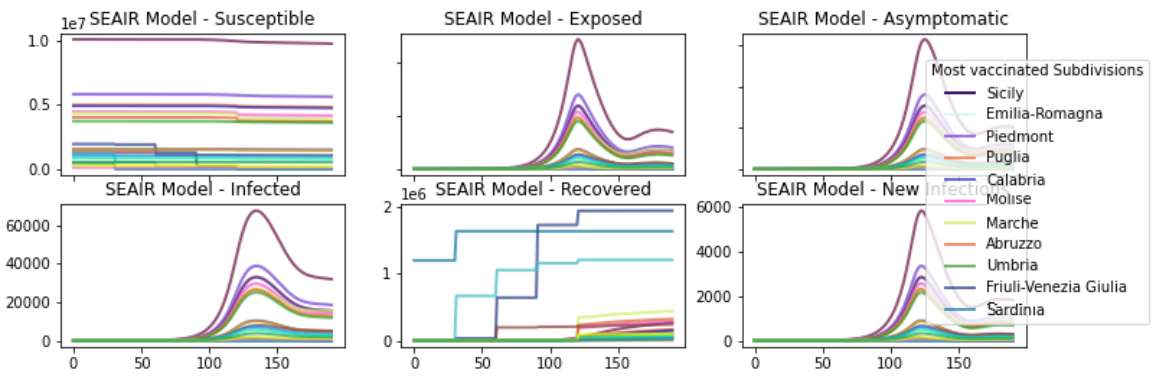


Figure 8.38: Plots of Susceptible, Exposed, Asymptomatic, Infected, Recovered and New infections per day for each Region of the SEAIR Model with the best vaccination plan found.

The comparison table between values obtained with and without vaccination are shown in table 8.15. The first column of the table corresponds to the evaluated criterion that can be the total infected population, maximum quantity of infected and asymptomatic at the same time, maximum quantity of infected at the same time and maximum new infected per day (or the peak of the curve). The next column show the corresponding results with the best vaccination applied and without vaccination. The last column shows the improving percentage after the vaccination plan was applied.

This vaccination plan covers about a 9.94% of the total Italy population and it is observed that, in this case, the decrease percentages are greater to this value. The criterion that most decrease after the vaccination plan is applied is the total infected population with a considerable value of 34.565%. These differences are not clear in plots, but this is because the numbers obtained for Italy population are in a larger scale due to their population. The decrease percentage of maximum infected, maximum asymptomatic and maximum new infected per day are close between them with percentages of 19.505% and 19.809%.

	With Best Vaccination	Without Vaccination	Decrease Percentage
Total Infected	1,447,971.778	2,212,845.957	34.565%
Maximum Infected and Asymptomatic	441,802.872	548,857.096	19.505%
Maximum Infected	369,804.099	461,722.174	19.908%
Maximum new infected per day	32,016.950	39,460.132	18.863%

Table 8.15: Comparison table of infections with and without vaccination.

8.3.7 UK

First Wave

In order to apply the SEAIRV model, and adjustment between curves was done. To get a similar curve for this first wave with a peak value on approximately day 60 it was necessary a basic reproduction number (R_0) with a high value of 5.5. This probably due to the low number of COVID-19 tests made at the beginning of the outbreak.

In figures 8.39 and 8.40 it can be observed the resulting curve of new infected per day and cumulative cases respectively, after the adjustments to the data of COVID-19 test of United Kingdom. The x-axis shows the days and y-axis show the number of individuals.

It is observed that the curve more than a peak has a constant high value for about 20 days. Also it is noticed that the curve decreases slowly. From the cumulative cases it is observed that both curves are really similar, also at the beginning of the COVID-19 tests curve from day 0 to 40 the uncertainty of the number of tests made it is noticeable.

Second Wave

The result of the adjustment between curves of the second wave is shown in figures 8.41 and 8.42. The x-axis shows the days since the start of the disease and the y-axis shows the number of individuals.

In this case a vacation period was seat between days 190 and 220 with an increase of the movement of infected and non infected population of 0.0025.

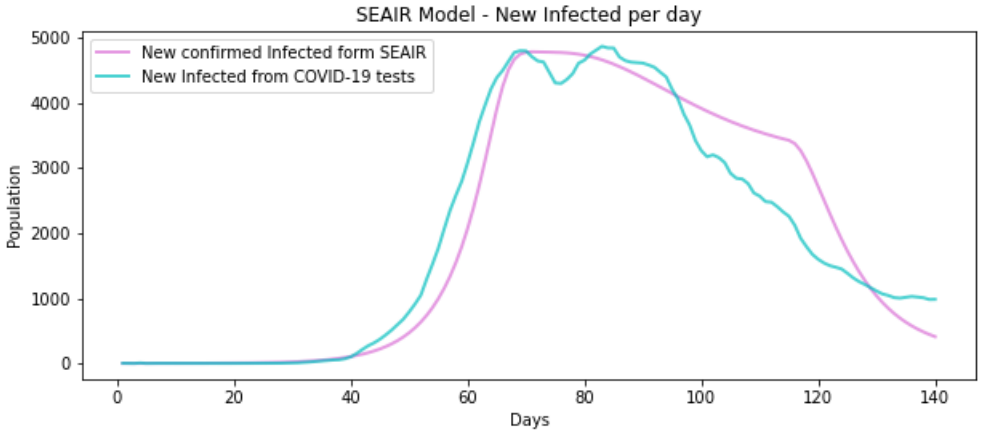


Figure 8.39: Comparison of new infected people between SEAIR Model and COVID-19 tests in United Kingdom. First wave.

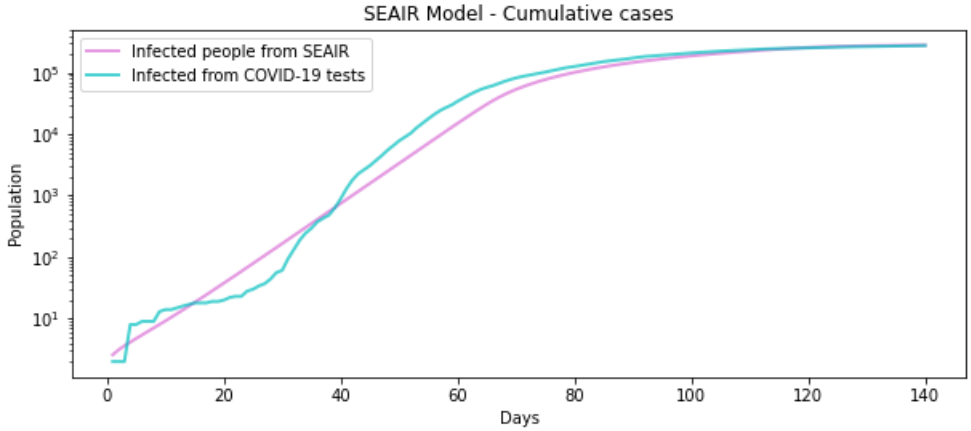


Figure 8.40: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in United Kingdom. First wave.

The obtained new infected per day curve is quite peculiar, it has two peaks but the first one is more close to an equilibrium state. From the cumulative cases it can be noticed a slow increase of the cases between day 150 and 200 because of the decay in the number of new infected population.

From the proposal execution, the vaccination plan was conformed of 1,200,000 vaccines per period. Part of it is shown in table 8.16. This table shows the six subgroups to which the highest number of vaccines were allocated.

The first column of the table shows the period of vaccination, all other columns show the subgroups considered. In this case six of the twelve regions of United Kingdom: South East, East of England, West Midlands, South West, East Midlands

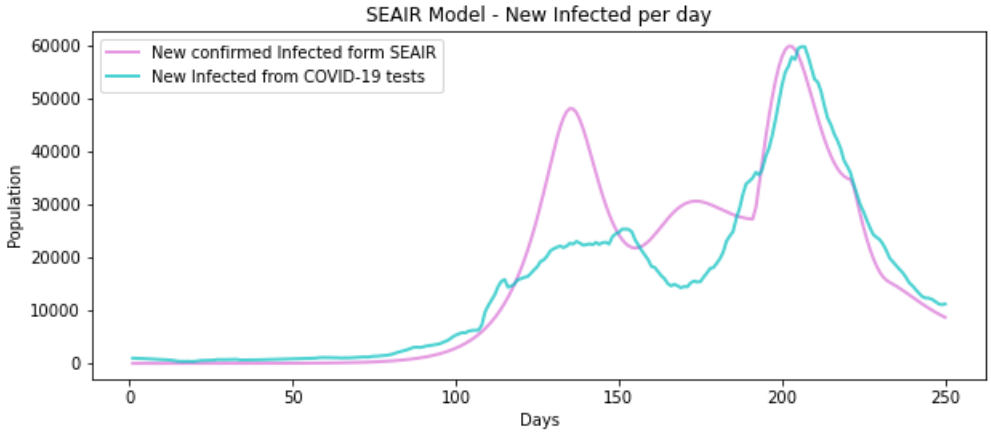


Figure 8.41: Comparison of new infected people between SEAIR Model and COVID-19 tests in United Kingdom. Second wave.

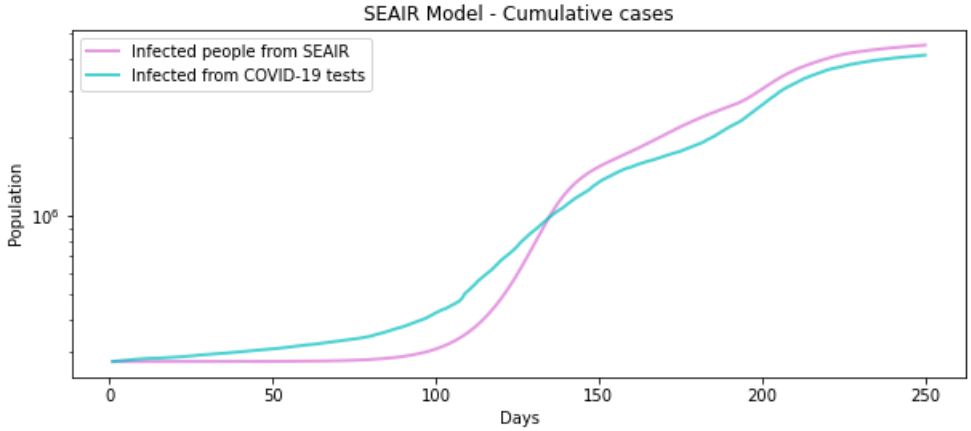


Figure 8.42: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in United Kingdom. Second wave.

and Northern Ireland.

It is observed that in four of the five vaccination periods all the vaccines were allocated to one subgroup. Despite of this, three different subgroups were allocated. In this case, again the vaccines are not allocated into the subgroups with the highest number of individuals.

The resulting curve after the vaccination process is shown in figures 8.43 and 8.44. Here x-axis shows the days and y-axis show the number of individuals.

It can be noticed from the new infected per day curve that the second wave or second peak decreases considerably. Also, that the curve perform a small shift to the right. From the cumulative cases, the curve starts increasing a bit after

	South East	East of England	West Midlands	South West	East Midlands	Northern Ireland
1^{st} Vaccination	0	1,200,000	0	0	0	0
2^{nd} Vaccination	0	0	0	0	0	1,200,000
3^{rd} Vaccination	15	0	508,012	19,259	1,210	671,501
4^{th} Vaccination	0	0	1,200,000	0	0	0
5^{th} Vaccination	0	0	1,200,000	0	0	0

Table 8.16: Vaccines corresponding to the most vaccinated subdivision of the best vaccination plan obtained from SEAIRV Model.

compared to the curve without vaccination.

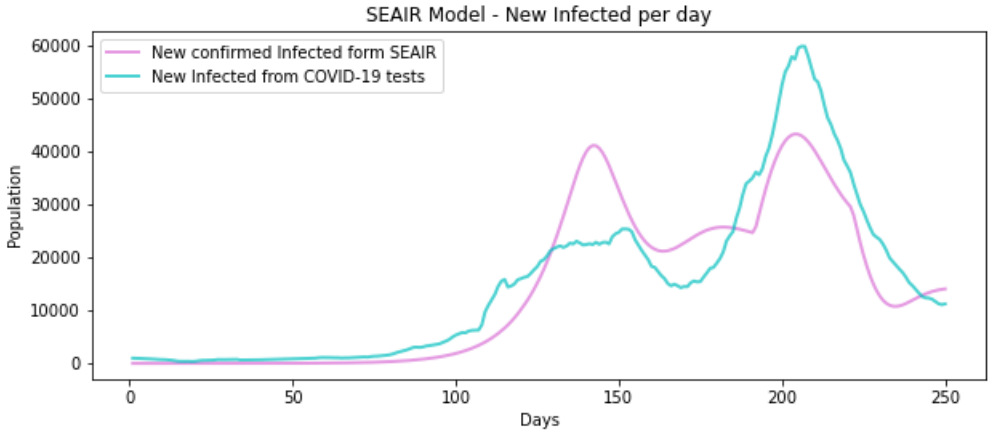


Figure 8.43: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in United Kingdom second wave after applying the best vaccination plan found.

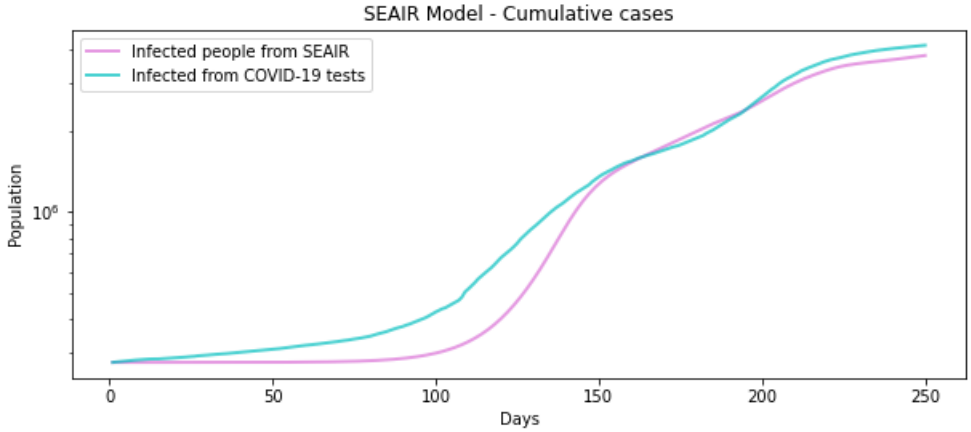


Figure 8.44: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in United Kingdom second wave after applying the best vaccination plan found.

Figure 8.45 shows the curves of each of the subgroups in United Kingdom when

the best vaccination plan found is applied. In these figures, the number of susceptible (S), exposed (E), asymptomatic (A), infected (I), recovered (R) and new infections overtime are presented. In all these plots the x-axis shows the days since the start of the disease and the y-axis shows the number of individuals.

From the susceptible plot, it is noticeable that the vaccinations were made to subgroups with medium or low total population than other possible subgroups. This may be due to their location and their contact with another subgroups. From the infected curves it can be noticed how West Midlands has only one wave or one peak value and its curve grows slowly.

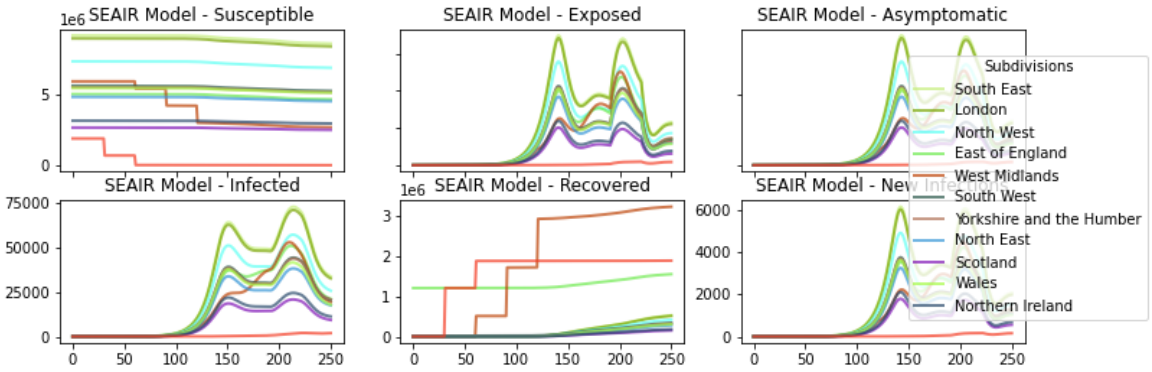


Figure 8.45: Plots of Susceptible, Exposed, Asymptomatic, Infected, Recovered and New infections per day for each Region of the SEAIR Model with the best vaccination plan found.

The results of the vaccination plan compared to not vaccination are shown in table 8.17.

The first column of the table corresponds to the evaluated criterion: the total infected population, maximum number of infected and asymptomatic at the same time, maximum number of infected at the same time and maximum new infected per day (or the peak of the curve). Next columns show the corresponding result with the best vaccination plan, without vaccination and the corresponding perceptual gain.

This vaccination plan covers about a 9.48% of the total United Kingdom population and it is observed that, in this case, the decrease percentages are greater than this value. Unlike previous instances, the decrease percentage of total infected is the lowest from the four criterion. This may be due to a decrease of the peak of the curve can lead to more movement of the population and also a more flatten and open curve. The values of maximum infected and asymptomatic and max-

imum infected have a similar percentage. The most reduced criterion with this vaccination plan was the maximum new infected per day (the peak of the curve).

	With Best Vaccination	Without Vaccination	Decrease Percentage
Total Infected	3,516,434.909	4,240,838.973	17.082%
Maximum Infected and Asymptomatic	577,270.889	735,925.857	21.559%
Maximum Infected	519651.316	659,949.229	21.259%
Maximum new infected per day	43,277.922	59,895.644	27.744%

Table 8.17: Comparison table of infections with and without vaccination.

8.3.8 Chile Metropolitan Region

First Wave

In order to apply the SEAIRV model, and adjustment between curves was done. For Chile Metropolitan Region, the model was adjusted to only the first wave. The curve of new infected per day after these adjustments is presented in figure 8.46 and the cumulative infected population in figure 8.47. For these both graphs the x-axis shows the period of time (days) and the y-axis shows the number of individuals.

In this case it is observed that the peak value is located between day 100 and 125. The adjusted SEAIR curve and the COVID-19 tests have a similar shape with respect to the new infected per day curve. From the cumulative cases curve, it is obtained that the values after day 100 are very similar.

The vaccination plan obtain was conformed of 250,000 vaccines per period. Part of it is showed in table 8.18. This table shows the six subgroups to which the highest number of vaccines were allocated.

The first column of the table shows the period of vaccination, all other columns show the subgroups considered. In this case six of the fifty-two districts of the Chile Metropolitan Region: Puente Alto, Santiago, San Bernardo, Las Condes, Providencia and Huechuraba.

In this case as much as a subgroup with a high quantity of population and high contact with another subgroups as Puente Alto is allocated with more vaccines.

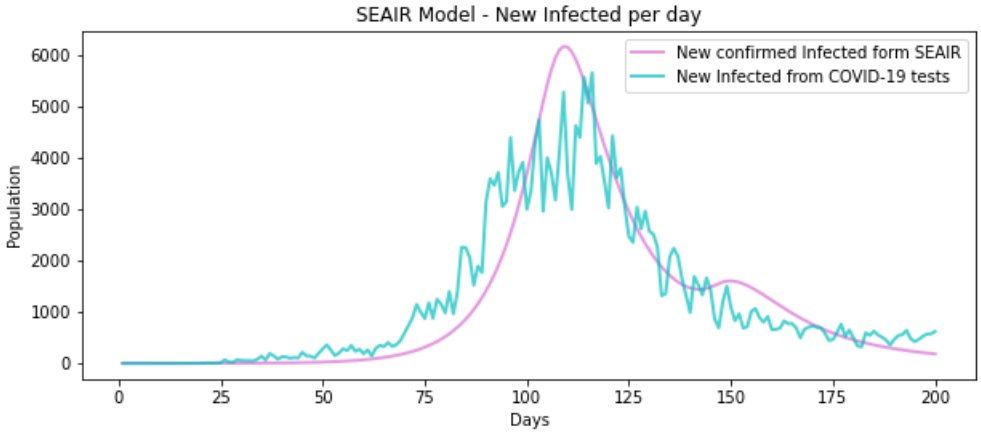


Figure 8.46: Comparison of new infected people between SEAIR Model and COVID-19 tests in Chile Metropolitan Region. First wave.

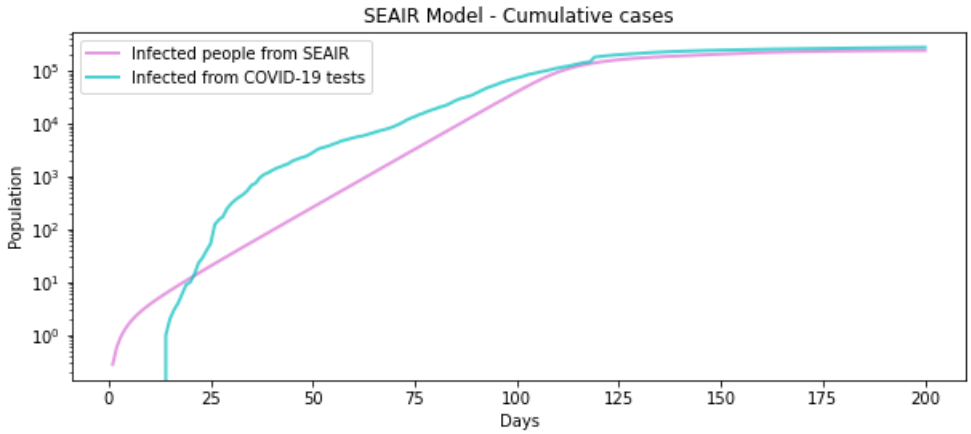


Figure 8.47: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Chile Metropolitan Region. First wave.

Also, two of the most visited subgroups (Santiago and Providencia) are in the group of the ones with more vaccines.

	Puente Alto	Santiago	San Bernardo	Las Condes	Providencia	Huechuraba
1^{st} Vaccination	0	0	250,000	0	0	0
2^{nd} Vaccination	93,336	22,008	0	22,284	108,346	0
3^{rd} Vaccination	0	86,551	0	56,088	24,212	19,110
4^{th} Vaccination	0	250,000	0	0	0	0
5^{th} Vaccination	59,053	46,278	47,977	141	7,004	57,836

Table 8.18: Vaccines corresponding to the most vaccinated subdivision of the best vaccination plan obtained from SEAIRV Model.

The resulting curves after the vaccination process is shown in figures 8.48 and 8.49. Here x-axis shows the days and y-axis show the number of individuals.

From the new infected per day curve it can be observed that the peak decreases considerably, also that the curve shift to the right and the decay after the peak goes down more abruptly. From the cumulative cases a slight decrease is observed.

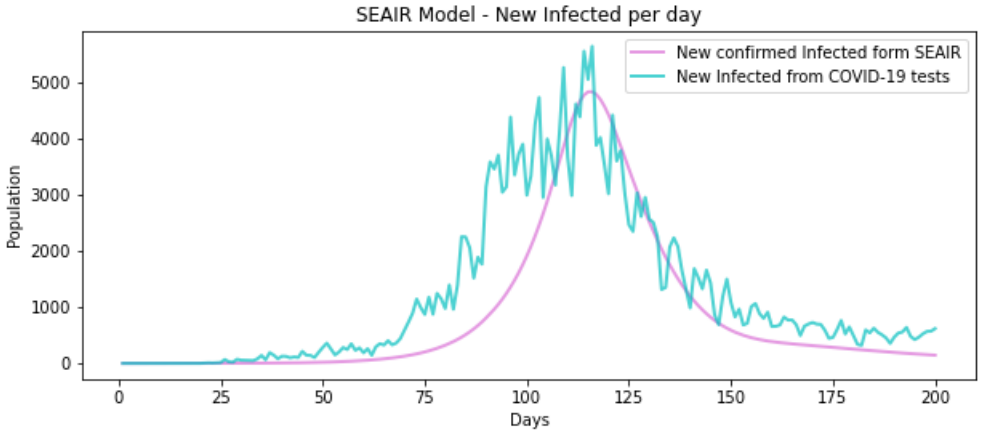


Figure 8.48: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Chile Metropolitan Region after applying the best vaccination plan found.

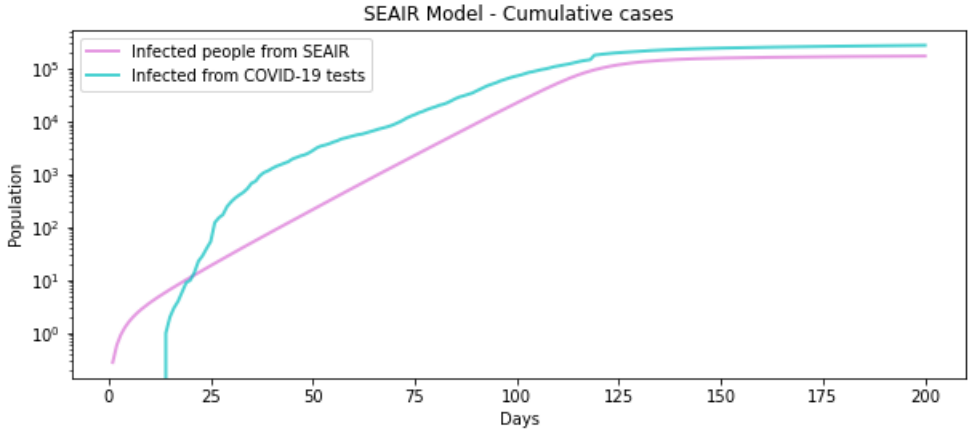


Figure 8.49: Comparison of cumulative infected people between SEAIR Model and COVID-19 tests in Chile Metropolitan Region after applying the best vaccination plan found.

Figure 8.50 shows the curves of each of the subdivisions in Chile Metropolitan Region when the best vaccination plan found was applied. In these figures, the number of susceptible (S), exposed (E), asymptomatic (A), infected (I), recovered (R) and new infections overtime are presented. In all these plots the x-axis shows the period of time (in days) and the y-axis shows the number of individuals.

From the susceptible plot a variety between the population of each subgroup vaccinated is noticed. Also, between the asymptomatic and infected curves it can

be observed how the asymptomatic curve comes downs after the peak value faster than the infected curve does.

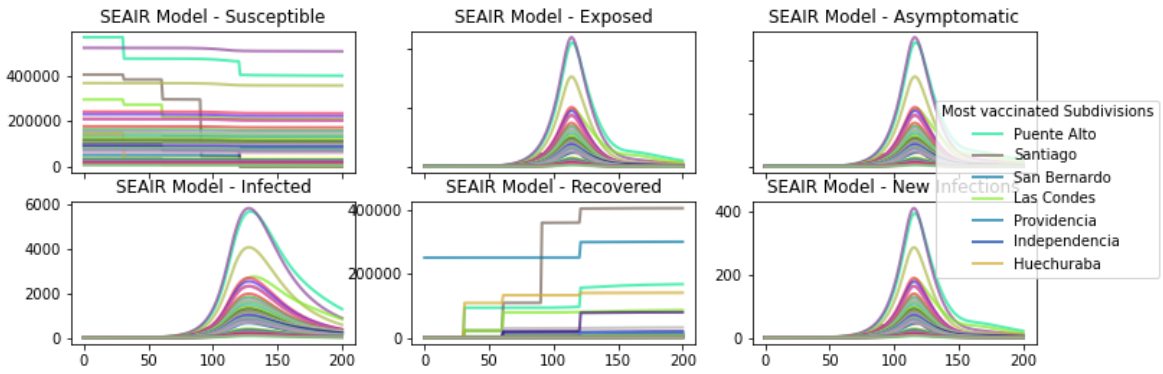


Figure 8.50: Plots of Susceptible, Exposed, Asymptomatic, Infected, Recovered and New infections per day for each Region of the SEAIR Model with the best vaccination plan found.

The results of the vaccination plan compared to the not vaccination scenario are shown in table 8.19. The first column of the table corresponds to the evaluated criterion: the total infected population, maximum number of infected and asymptomatic at the same time, maximum number of infected at the same time and maximum new infected per day (or the peak of the curve). Next columns show the corresponding result with the best vaccination plan, without vaccination and the corresponding perceptual gain.

This vaccination plan covers about a 16.35% of the total Chile Metropolitan Region population and it is observed that, in this case, the decrease percentages are close to this value with the exception of the total infected percentage that reaches to 27.62%. In this case for the total infected and new infected per day the decrease values are high, this may be due to the movement restrictions in this case where a diminution of the new infected per day does not influences enough the movement to increase the total infected population.

8.4 Chapter Summary

In this chapter an analysis and conclusion for the experiments was made. Firstly, the SEAIRV parameter setting experiments results were presented. From the movements relevance it was obtained that for three of the four instances tested the best

	With Best Vaccination	Without Vaccination	Decrease Percentage
Total Infected	177,132.169	245,888.633	27.962%
Maximum Infected and Asymptomatic	73,278.402	89,129.215	17.784%
Maximum Infected	69,561.196	84,617.419	17.793%
Maximum new infected per day	4,838.192	6,160.004	21.458%

Table 8.19: Comparison table of infections with and without vaccination.

movement was to just use the movement *Give Random* but due to the existence of other type of instances that need completely the opposite it was decided to use probabilities 30% to *Give Random*, 10% of *Swap Random* and 10% to *Invert random*. From the initialization procedures it has been obtained that *All to One* gives the best results for three of the instances and it set as the initialization method of the algorithm. Finally, from the objective functions it is obtained that influence in the selection of a solution, is possible to obtain solutions with highest value of maximum infected and lowest value of total infected, for this the percentage of each function was seated to 50% and 50%. From the vaccination plan obtained per each problem instance graphs with and without vaccination were shown, explaining the differences and changes to the COVID-19 test data curves. Furthermore, the most vaccinated subgroups were shown and the behavior of each subgroup after vaccination. Finally, for each instance the evaluated criterion were shown and explained.

Conclusions

The exposed document focus principally on modelling and simulating the COVID-19 spread through subgroups as closet as possible to the observed behaviour. This mainly considering geographic subdivisions and interaction, and with the idea of finding efficient vaccination plans considering interval of periods where vaccines are available, and with the aim of decreasing the maximum number of infected people at the same time and the total infected population of the obtained model.

The Susceptible-Exposed-Asymptomatic-Infected-Recovered model is used here to simulate the disease spread. This model is based on susceptible population that get exposed implying that these individuals are on a incubation state of the virus. After an amount of time expressed as incubation rate (δ), they get infectious (can infect susceptible population). The infectious population was divided with a certain proportion (μ) into asymptomatic individuals without symptoms and infected individuals with symptoms. This because different movement restrictions are applied to them. Asymptomatic individuals can be tested and discovered as infectious getting from asymptomatic to infected individuals with a discovery rate (η). Infectious population get recovered after a certain time expressed as the recovery rate (γ).

Eight cases from seven countries with high COVID-19 test performed, variety of subdivisions and population sizes were selected for experiments. These countries were Denmark, Qatar, Austria, Belgium, United Kingdom, Chile and Italy. From these cases, each input file was created based on simple geographical data, population data and new infected and cumulative cases of their COVID-19 tests. The procedure to construct each problem instance consists of a set of the population and subgroups from real data, the creation of a contact matrix using a two step procedure that first creates a random graph based on how many connections or

how important is each subgroup and, secondly it performs a geographical adjustment based on the proximity of the subdivisions for seven of the eight cases. For the other case, this was based on some information of the movement of population when they travel from home to work and schools.

The adjustments made to get closer to each COVID-19 test curve were based on COVID-19 parameters as incubation rate (δ), recovery rate(γ) and percentage of infected that are asymptomatic (μ), basic reproduction number R_0 and the detection of asymptomatic (η) and SEAIR parameters that determines the restriction movement for infected population and non infected population. It was sought to assimilate as much as possible the peak of the new infected curve, the day of the peak, the inclination in cumulative cases curve and the shape of the fall after the curve.

A Susceptible-Exposed-Asymptomatic-Infected-Recovered Vaccination (SEAIRV) local search based algorithm was implemented to determine an efficient vaccination distribution per period or also called vaccination plan. The algorithm uses a tabu search scheme for each period of the vaccination process. Five initialization options were implemented: *inner interaction*, *outer interaction*, *mixed interaction*, *equity* and *all to one*. Moreover, it implements three movements: *give random*, *swap random* and *invert random* along with a tabu list that prevent cycles during the search process. Two main objectives were considered: the minimization of the maximum number of infected at the same time and the minimization of the total infected population obtained from the SEAIR model. The algorithm implements an evaluation function that uses a weighted sum approach.

Two types of experiments were performed. The first ones were oriented to analyze the SEAIRV parameter values and the second ones were oriented to get the vaccination plans for each problem instance. From the first experiments it was possible to conclude that the best parameters values for the SEAIRV algorithm were the *all to one* initialization due to the lower time obtained with it with presumed faster convergence and stagnation. Probabilities of 80%, 10% and 10% for the movements *give random*, *swap random* and *invert random* respectively being that there is a big necessity of the movement *give random* in a way to change proportion of the number of vaccines per subgroup and the other two movements can support to a faster convergence. Finally, a distribution of the half of importance to each of the objective functions because of the time taken to converge and to search for solutions with both objectives low.

Another of the conclusions obtained with these experiments were that there are problem instances were these objectives are opposite, i. e. there are solutions with more infected population at the same time but lower value of total infected in the period.

In the second set of experiments, the SEAIRV algorithm was run and the best vaccination plan obtained was applied and analyzed. Furthermore, the curves with and without vaccination, the values of total infected, total infected and asymptomatic, new infected per day and vaccines per subdivision were shown. From this experiments it could be concluded that for the proposed method when looking for the combination of vaccines per period, taking into account that period and not future ones, the first infected population influence in most of the cases to allocate vaccines to these subgroups. Furthermore, it has been shown that despite of the percentage of the population vaccinated with small percentages it is possible to make greater changes depending on the vaccination plan, the reproduction number and the movements of the infected and not infected population.

As future work, there are some scenarios that can be studied. To use real movement data between subdivisions to get problem instances closer to reality. Also, to add more subgroups with different ages and prioritize to vaccinate some of them or that they have different movement or contact. Change the way to find vaccines per period, focusing on a time after the last vaccination to see if find better combinations. Find new movements to allow a faster convergence to high quality solutions.

Bibliography

[1] Nicolas Bacaër. "McKendrick and Kermack on epidemic modelling (1926–1927)". In: *A Short History of Mathematical Population Dynamics*. London: Springer London, 2011, pp. 89–96. ISBN: 978-0-85729-115-8. DOI: 10.1007/978-0-85729-115-8_16. URL: https://doi.org/10.1007/978-0-85729-115-8_16.

[2] Kirtikumar C. Badgujar, Vivek C. Badgujar, and Shamkant B. Badgujar. "Vaccine development against coronavirus (2003 to present): An overview, recent advances, current scenario, opportunities and challenges". In: *Diabetes Metabolic Syndrome: Clinical Research Reviews* 14.5 (2020), pp. 1361–1376. ISSN: 1871-4021. DOI: https://doi.org/10.1016/j.dsx.2020.07.022. URL: http://www.sciencedirect.com/science/article/pii/S1871402120302708.

[3] Carl Heneghan, Jon Brassey, Tom Jefferson. *CEBM portion of asymptomatic of covid-19*. URL: http://www.cebm.net/covid-19/covid-19-what-proportion-are-asymptomatic/.

[4] S. Chen, M. Small, and X. Fu. "Global stability of epidemic models with imperfect vaccination and quarantine on scale-free networks". In: *IEEE Transactions on Network Science and Engineering* (2019), pp. 1–1.

[5] CIAE - Universidad de Chile. *Desplazamiento de los estudiantes en Chile*. URL: http://www.ciae.uchile.cl/index.php?page=view_noticias&langSite=es&id=1885.

[6] Sebastián Contreras et al. "A multi-group SEIRA model for the spread of COVID-19 among heterogeneous populations". In: *Chaos, Solitons Fractals* 136 (2020), p. 109925. ISSN: 0960-0779. DOI: https://doi.org/10.1016/j.chaos.2020.109925. URL: http://www.sciencedirect.com/science/article/pii/S0960077920303246.

[7] Max Leo Correa Cordova, Abebe Geletu, and Pu Li. "Optimal Scheduling of Vaccination Campaigns Using a Direct Dynamic Optimization Method". In: *IFAC-PapersOnLine* 49.26 (2016). Foundations of Systems Biology in Engineering - FOSBE 2016, pp. 207–212. ISSN: 2405-8963. DOI: https://doi.org/10.1016/j.ifacol.2016.12.127. URL: http://www.sciencedirect.com/science/article/pii/S2405896316327902.

[8] Andrés R. [da Cruz], Rodrigo T.N. Cardoso, and Ricardo H.C. Takahashi. "Multiobjective synthesis of robust vaccination policies". In: *Applied Soft Computing* 50 (2017), pp. 34–47. ISSN: 1568-4946. DOI: https://doi.org/10.1016/j.asoc.2016.11.010. URL: http://www.sciencedirect.com/science/article/pii/S1568494616305798.

[9] DATA UC. *Visualizador Covid-19 Chile*. URL: https://coronavirus.mat.uc.cl/.

[10] Shakiba Enayati and Osman Y. Özaltın. "Optimal influenza vaccine distribution with equity". In: *European Journal of Operational Research* 283.2 (2020), pp. 714–725. ISSN: 0377-2217. DOI: https://doi.org/10.1016/j.ejor.2019.11.025. URL: http://www.sciencedirect.com/science/article/pii/S0377221719309361.

[11] Gabriel Fabricius and Alberto Maltz. "Exploring the threshold of epidemic spreading for a stochastic SIR model with local and global contacts". In: *Physica A: Statistical Mechanics and its Applications* 540 (2020), p. 123208. ISSN: 0378-4371. DOI: https://doi.org/10.1016/j.physa.2019.123208. URL: https://www.sciencedirect.com/science/article/pii/S0378437119318035.

[12] Libi Fu et al. "Simulation of emotional contagion using modified SIR model: A cellular automaton approach". In: *Physica A: Statistical Mechanics and its Applications* 405 (2014), pp. 380–391. ISSN: 0378-4371. DOI: https://doi.org/10.1016/j.physa.2014.03.043. URL: https://www.sciencedirect.com/science/article/pii/S0378437114002386.

[13] Yong Gao. "Treewidth of Erdös–Rényi random graphs, random intersection graphs, and scale-free random graphs". In: *Discrete Applied Mathematics* 160.4 (2012), pp. 566–578. ISSN: 0166-218X. DOI: https://doi.org/10.1016/j.dam.2011.10.013. URL: https://www.sciencedirect.com/science/article/pii/S0166218X11003726.

[14] Global Change Data Lab. *Ourworldindata*. URL: https://ourworldindata.org.

[15] Gobierno Regional Metropolitano de Santiago. *Diagnostico de la Region Metropolitana de Santiago para la Estrategia Regional de Desarrollo 2012*. URL: https://www.gobiernosantiago.cl/wp-content/uploads/2014/doc/estudios/Diagnostico_de_la_Region_Metropolitana_de_Santiago_para_la_Estrategia_Regional_de_Desarrollo,_2012.pdf.

[16] M. Halloran, Claudio Struchiner, and I. Longini. "Study Designs for Evaluating Different Efficacy and Effectiveness Aspects of Vaccines". In: *American journal of epidemiology* 146 (Dec. 1997), pp. 789–803. DOI: 10.1093/oxfordjournals.aje.a009196.

[17] Donghyun Kim et al. "On efficient vaccine distribution strategy to suppress pandemic using social relation". In: *Discrete Mathematics, Algorithms and Applications* 08.01 (2016), p. 1650010. DOI: 10.1142/S1793830916500105. eprint: https://doi.org/10.1142/S1793830916500105. URL: https://doi.org/10.1142/S1793830916500105.

[18] Michael Y. Li and James S. Muldowney. "Global stability for the SEIR model in epidemiology". In: *Mathematical Biosciences* 125.2 (1995), pp. 155–164. ISSN: 0025-5564. DOI: https://doi.org/10.1016/0025-5564(95)92756-5. URL: http://www.sciencedirect.com/science/article/pii/0025556495927565.

[19] Yen-Chin Liu, Rei-Lin Kuo, and Shin-Ru Shih. "COVID-19: The first documented coronavirus pandemic in history". In: *Biomedical Journal* (2020). ISSN: 2319-4170. DOI: https://doi.org/10.1016/j.bj.2020.04.007. URL: http://www.sciencedirect.com/science/article/pii/S2319417020300445.

[20] Alberto Maltz and Gabriel Fabricius. "SIR model with local and global infective contacts: A deterministic approach and applications". In: *Theoretical Population Biology* 112 (2016), pp. 70–79. ISSN: 0040-5809. DOI: https://doi.org/10.1016/j.tpb.2016.08.003. URL: https://www.sciencedirect.com/science/article/pii/S0040580916300454.

[21] Bimal Kumar Mishra et al. "COVID-19 created chaos across the globe: Three novel quarantine epidemic models". In: *Chaos, Solitons Fractals* 138 (2020), p. 109928. ISSN: 0960-0779. DOI: https://doi.org/10.1016/j.chaos.2020.109928. URL: http://www.sciencedirect.com/science/article/pii/S0960077920303271.

[22] National Center for Immunization and Respiratory Diseases. *CDC coronavirus planning scenarios*. URL: www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html.

[23] C.T. Ng et al. "A multi-criterion approach to optimal vaccination planning: Method and solution". In: *Computers Industrial Engineering* 126 (Oct. 2018). DOI: 10.1016/j.cie.2018.10.018.

[24] Pauta Chile. *Donde viven y donde trabajan*. URL: https://www.pauta.cl/ciudad/mapa-de-residencia-de-los-trabajadores-en-la-region-metropolitana-santiago.

[25] G. Schneckenreither et al. "Modelling SIR-type epidemics by ODEs, PDEs, difference equations and cellular automata - A comparative study". In: *Simulation Modelling Practice and Theory* 16.8 (2008). EUROSIM 2007, pp. 1014–1023. ISSN: 1569-190X. DOI: https://doi.org/10.1016/j.simpat.2008.05.015. URL: https://www.sciencedirect.com/science/article/pii/S1569190X08001160.

[26] Matthew W. Tanner, Lisa Sattenspiel, and Lewis Ntaimo. "Finding optimal vaccination strategies under parameter uncertainty using stochastic programming". In: *Mathematical Biosciences* 215.2 (2008), pp. 144–151. ISSN: 0025-5564. DOI: https://doi.org/10.1016/j.mbs.2008.07.006. URL: http://www.sciencedirect.com/science/article/pii/S0025556408001156.

[27] George G. Vega Yon, Andrew Slaughter, and Kayla de la Haye. "Exponential random graph models for little networks". In: *Social Networks* 64 (2021), pp. 225–238. ISSN: 0378-8733. DOI: https://doi.org/10.1016/j.socnet.2020.07.005. URL: https://www.sciencedirect.com/science/article/pii/S0378873320300496.

[28] D Volchenkov and Ph Blanchard. "An algorithm generating random graphs with power law degree distributions". In: *Physica A: Statistical Mechanics and its Applications* 315.3 (2002), pp. 677–690. ISSN: 0378-4371. DOI: https://doi.org/10.1016/S0378-4371(02)01004-X. URL: https://www.sciencedirect.com/science/article/pii/S037843710201004X.

[29] Xinwei Wang et al. "Optimal vaccination strategy of a constrained time-varying SEIR epidemic model". In: *Communications in Nonlinear Science and Numerical Simulation* 67 (2019), pp. 37–48. ISSN: 1007-5704. DOI: https://doi.org/10.1016/j.cnsns.2018.07.003. URL: http://www.sciencedirect.com/science/article/pii/S100757041830217X.

[30] Qingchu Wu and Yijun Lou. "Local immunization program for susceptible-infected-recovered network epidemic model". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 26.2 (2016), p. 023108. DOI: 10.1063/1.4941670. eprint: https://doi.org/10.1063/1.4941670. URL: https://doi.org/10.1063/1.4941670.

[31] Li-Meng Yan et al. "Combined use of live-attenuated and inactivated influenza vaccines to enhance heterosubtypic protection". In: *Virology* 525 (2018), pp. 73–82. ISSN: 0042-6822. DOI: https://doi.org/10.1016/j.virol.2018.09.007. URL: http://www.sciencedirect.com/science/article/pii/S0042682218302782.