



A simple but complex enough θ -SIR type model to be used with COVID-19 real data. Application to the case of Italy

A.M. Ramos^a, M.R. Ferrández^b, M. Vela-Pérez^a, A.B. Kubik^a, B. Ivorra^a

^a MOMAT Research Group
Interdisciplinary Mathematics Institute
Complutense University of Madrid

^b Department of Computer Science,
University of Almería

emails: angel@mat.ucm.es; mrferrandez@ual.es; maria.vela@ucm.es; akubik@ucm.es; ivorra@ucm.es

23 December 2020
(update of a previous version published on 23 July 2020)

DOI link: <http://www.doi.org/10.13140/RG.2.2.32466.17601>

Accepted for publication in *Physica D: Nonlinear Phenomena*

DOI link: <https://doi.org/10.1016/j.physd.2020.132839>¹

Abstract

Since the start of the COVID-19 pandemic in China many models have appeared in the literature, trying to simulate its dynamics. Focusing on modeling the biological and sociological mechanisms which influence the disease spread, the basic reference example is the SIR model. However, it is too simple to be able to model those mechanisms (including the three main type of control measures: social distancing, contact tracing and health system measures) to fit real data and to simulate possible future scenarios. A question, then, arises: how much and how do we need to complexify a SIR model? We develop a θ -SEIHRD model, which may be the simplest one satisfying the mentioned requirements for arbitrary territories and can be simplified in particular cases. We show its very good performance in the Italian case and study different future scenarios.

Keywords: Mathematical model, θ -SEIHRD model, COVID-19, Coronavirus, SARS-CoV-2, Pandemic, Numerical simulation, Parameter estimation, basic reproduction number, effective reproduction number

¹Citation: A.M Ramos, M.R. Ferrández, M. Vela-Pérez, A.B. Kubik and B. Ivorra. A simple but complex enough θ -SIR type model to be used with COVID-19 real data. Application to the case of Italy. *Physica D: Nonlinear Phenomena*. Available online (132839). 2021. DOI: <https://doi.org/10.1016/j.physd.2020.132839>

Data and code availability

All the data and all the model simulation and figure codes used in this manuscript are available at

<https://www.ucm.es/momat/software-momat> and at <https://github.com/momat-ucm/T-SIR-T>.

The core model was written in Java and the figures were obtained using MATLAB.

1 Introduction

The outbreak of novel coronavirus disease 2019 (COVID-19) continues to spread rapidly around the world [60]. Originated in Wuhan in central China, where it was already active at the beginning of December 2019, reached multiple countries in merely a month until it became a pandemic as of 11 March 2020 [63] when the disease was confirmed in more than 118,000 cases reported globally in 114 countries.

Although some authors (see, e.g., [37]) propose SEIR type models with little variations, COVID-19 is a disease caused by a new virus and needs a model taking into account its known specific characteristics. In particular, it would be convenient to develop a model which satisfies the following: a) includes the biological and sociological mechanisms influencing the disease spread, b) its parameters have a clear interpretation in terms of those mechanisms (such as a contact rate, fatality ratios, duration of infection, etc.), and c) incorporates

- the effect of undetected infected people (see [39]), being able to show the dependence of the impact of COVID-19 on the percentage of detected cases over the real total infected cases,
- the effect of different infectiousness conditions of hospitalized people (differentiating those with mild and severe conditions that will recover from those who will finally die),
- the estimation of the needs of beds in hospitals (which is one of the major problems for policy makers addressing COVID-19),
- the effect of different control measures.

Some of the models that appeared at the beginning of the pandemic, especially for the study of the first wave (see [19, 33, 37, 39]), do not incorporate all the special characteristics (undetected cases, undetected deaths, quarantine, social distancing, etc.) of the COVID-19 pandemic. Some of them are very complex in terms of the number of compartments and connections between them, trying to simulate many details that could probably be neglected. Furthermore, after developing a model, it would be desirable to check how it fits real data and to use it in order to study future scenarios with different control measures (here, we consider social distancing, contact tracing and health system measures).

This paper does not take into account the possibility of resusceptibility as it is done, for example, in [45]. This is due to the fact that the number of confirmed reinfections is still not relevant enough compared with the total figures of the pandemic. According to [48], only 27 cases of reinfection were documented worldwide when it was written. Furthermore, the webpage <https://bnonews.com/index.php/2020/08/covid-19-reinfection-tracker/> reports an updated list of reinfections, with a total of 31 cases, at the time we wrote this article.

In [20], a disease contact rate (β) and R_0 with different values before and after lockdown of countries are considered and applied in a SEIR type model for the confirmed cases tested with several countries. The model presented here allows to model not only confirmed cases but also deaths, hospitalized and recovered by considering different values of β that depend on the particular control measures for every time period.

Concerning the particular case of Italy, some studies have dealt with its particular situation, such as [49], [10], [18] and [4]. It is important to note that all the models presented in these works only serve the particular case of Italy, and are not general models applicable to other countries.

The paper [49] is an initial work published in mid-March 2020, which highlights the importance of this pandemic and the serious consequences it can have in countries like Italy where data show rapid progress in infections, which can cause a health system collapse due to the lack of beds in hospitals and ICU posts. The data are fitted using

an exponential type model, which is not realistic for long-term modeling, although it allows a wake-up call to the country's health authorities.

In [10], the distribution in terms of mobility is also studied through a network of the different Italian regions. It is confirmed the effectiveness at a regional level of the national lockdown strategy and proposed coordinated regional interventions to prevent future national lockdowns. However, asymptomatic patients are not considered.

In [18], a geographical model is presented to describe the number of daily cases, hospitalized and deaths within the Italian regions for the first month of the disease and study the effects of drastic measures for transmission containment. Authors concentrate on estimating the effects of severe progressive restrictions posed to human mobility and human-to-human contacts in Italy. As a novelty, a presymptom compartment is considered. The proposed model is extended and further developed in [4], where the effect of mobility between regions and provinces within Italy is also considered, taking into account the control measures at all times to describe the variable of new hospitalizations. The number of people who have to be isolated in each region is calculated to avoid a drastic growth in the number of infected. This model is very interesting, because it incorporates the variable of mobility between regions, but only the number of new hospitalizations is described, without studying the number of deaths since they focus on the implications of the relaxation of confinement measures.

An interesting model appearing in [19], tested for the case of Italy, has some similarities with the one presented here, but there are some important differences to be remarked. First of all, the model we propose is a general one that may be used with any country or region, just considering the correspondent data, whereas the model in [19] seems to be defined for the particular case of the first wave in Italy, although it could be adapted to other territories. Secondly, our model considers the compartment for people in quarantine (Q , detected asymptomatic or with mild symptoms, infectious people) that is not the same as the proposed compartment for diagnosed (D , asymptomatic infected, detected) in the former model, which is even connected to the possibility of dying. Moreover, the general role of the undetected people in our model is completely different and some of them can even die (those who are in I_{D_u}).

In this paper, we develop a general deterministic epidemiological model able to simulate the biological and sociological COVID-19 mechanisms in different territories, to fit real data, and to simulate different future scenarios of control measures. The choice of using a deterministic model instead of a stochastic one is done as a first approach, since such kind of models present some advantages, such as: a low computational complexity allowing a better calibration of the model parameters or the possibility of using the theory of ordinary differential equations for suitably analyzing and interpreting the model. Furthermore, according to [12], deterministic models should be the first tool to be used when modeling a new problem. The authors of that work also note that the stochastic models are not suitable when it is difficult or impossible to determine the probability distribution, are difficult to analyze and require more data for the calibration of the model. We note that, although we use a deterministic model, we must deal with some probabilistic processes. In those cases, random values are substituted by their expected values.

We look for the simplest model, having the minimum complexity to take into account all the main non-negligible processes of this pandemic. When simulating the dynamics of the COVID-19 disease in a territory, our model can be used in its full general form or in simplified particular versions of it, depending on the way the territory deals with the pandemic. Of course, the model may be modified by introducing age structure, stochasticity, levels of uncertainty, mobility between territories, spatial structure (see, e.g., [35]), etc. in order to focus in special features of the pandemic or to refine some results.

The reference epidemiological model is the SIR one, which divides the population of a territory in three compartments: Susceptible (S), Infectious (I) and Recovered (R). For the COVID-19 pandemic these three compartments are not enough. First of all, the incubation period for the new coronavirus SARS-CoV-2 is between 5 and 6 days (see [64]), which is important to incorporate in the model by means of a compartment for Exposed (E) people. Secondly, estimation of needs of beds in hospital is important for policy makers, which implies the convenience of suitable compartments for people in Hospitals (H), following a compartment (I) for an initial stage of the infectious period. Moreover, the disease contact rates inside and outside hospitals are usually different. Thirdly, one of the main data reported each day is the number of Deaths (D), which deserves another compartment and makes convenient to split H in two branches, one for people in hospital that will recover (H_R) and the other one for people in hospital that will die (H_D). Therefore, we need, at least, a SEIHRD model.

Furthermore, one of the characteristics of COVID-19 is the existence of undetected infectious cases, either because they are asymptomatic (see [39]) or because, even if they have symptoms, the corresponding country/territory does

not have enough tests to check the infectiousness or the health system is saturated. In order to be able to consider the impact of this COVID-19 characteristic, two different branches should distinguish between detected and undetected cases and deaths. This can be done in terms of the instantaneous ratio $\theta(t)$ of new detected infectious cases per unit time over the total (detected and undetected) new infectious cases per unit time, generating the corresponding θ -SEIHRD model. As done previously with people in hospital, a distinction can be done between the undetected infectious people (compartment I_u) that will recover (compartment R_u) and the undetected infectious people (compartment I_{Du}) that will die (compartment D_u). Finally, some territories (like Italy) use a quarantine period (typically at home) for people leaving the hospital and for detected asymptomatic (or with mild symptoms) infectious people, which makes necessary a new compartment (Q), which completes the full θ -SEIHQRD general model. A diagram summarizing the model is given in Figure 1.

A detailed description and discussion of the model, its parameters and outputs is provided in Sections 2 and 3. We point out that all the parameters in the model have well-understood epidemiological interpretations such as average duration of incubation period, contact rates, etc. In Section 2, we also show some of the outputs of the model and how to compute them. Particular outputs of interest for epidemiologists are the computation of the associated basic reproduction number and the evolution of the effective reproduction number.

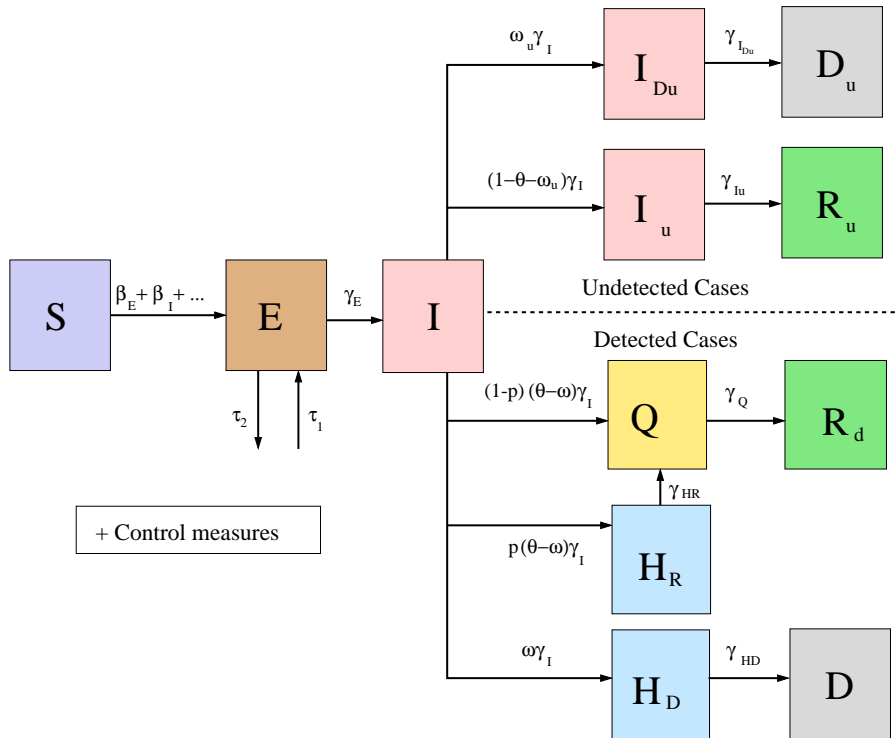


Figure 1: Diagram summarizing the model for COVID-19 given by system (1).

Once we have developed the model, the question is how to use it with real cases. Let us suppose we want to use it to study the dynamics of the COVID-19 in a particular territory. Three steps can be followed:

1. The first thing is to decide if all the compartments are necessary, or only some of them. For instance, in China and Senegal all infected people are treated in hospitals (or other kind of health facilities) and, therefore, compartment Q is not necessary [33]. However, other territories, like Italy or Spain, use the quarantine at home for infected people with very mild symptoms or not symptoms at all. Other possibilities of simplifying the model are also possible, for particular territories.

2. After deciding the compartments that will be used with a particular territory, the following step is to choose the parameters involved in the model. To do that we use values given in the literature (as the duration of the incubation period), parametrization of functions involved in the model and a multiobjective technique, to identify parameters. In Section 3 we explain how to do it. This is, actually, part of the validation process, consisting in comparing the results of the simulations (run with the chosen parameters) with the available data about the impact of the pandemic in the territory under study. Typically, a good fitting of simulations outputs to official temporal series (of cases, deaths, people in hospital, etc.) is necessary.
3. Finally, if the validation process is successful, we can simulate different future scenarios to study the efficiency of some control measures (being increased or relaxed), the impact of the undetected cases in the dynamics of the pandemic, etc.

In Section 4, we illustrate those three steps with the particular case of Italy.

2 Mathematical formulation of the model

In this section we give a detailed description of the θ -SEIHQRD model that we propose for the COVID-19. Then, we described some outputs of the model that will be used in the numerical experiments performed in Section 4.

2.1 The mathematical model

According to the known characteristics of the COVID-19 pandemic, we assume that each person is in one of the following compartments (see [2, 47, 61, 57]):

- Susceptible (denoted by S): The person is not infected by the disease pathogen.
- Exposed (denoted by E): The person is in the incubation period after being infected by the disease pathogen, and has no clinical signs. This stage is also called presymptomatic period. According to the available literature (see, e.g., [2]), there exist some evidences that a presymptomatic person could infect other people but with a lower probability than people in the infectious compartments. After the incubation period, the person passes to the Infectious compartment I .
- Infectious (denoted by I): After the incubation period, it is the first compartment of the infectious period, where nobody is expected to be detected yet. The person has finished the incubation period and may infect other people (even if he/she is asymptomatic, see, e.g., [61]). After this period, people in this compartment can be, either taken in charge by health authorities (and we classify them as hospitalized or quarantine, according to the severity of their symptoms), or not detected by authorities and continue as infectious (but in other compartment, I_u or I_{D_u}).
- Infectious but undetected that will survive the disease (denoted by I_u): After being in the compartment I , the person can still infect other people but is not detected and is not reported by authorities. We assume that only asymptomatic infected people or those with low or medium symptoms can reach this compartment, not the people who will die. After this period, people in this compartment survive the disease and pass to the Recovered compartment R_u .
- Infectious but undetected that will die (denoted by I_{D_u}): After being in the compartment I , the person can still infect other people, have clinical signs but not be detected and reported by authorities. We assume that a person in this state is part of a population that is highly vulnerable to the COVID-19 (such as people in retirement homes) and will die quickly without being treated in hospital. After being in this state, a person passes to the compartment D_u .

We remark that vulnerable people died outside hospitals during the first wave of the pandemic in some countries (such as Italy or Spain; see [32, 30, 24, 14, 41]). Indeed, due to the saturation of the whole health system (mainly

because of lack of beds and/or healthcare workers), this vulnerable population did not receive adequate medical cares and quickly died before being hospitalized. In some of those countries (such as Italy, see [32, 30]), a few a-posteriori estimations of the number of people dead due to COVID-19 and not detected have been reported by authorities or other organizations and, currently, efforts are made to avoid reproducing this extreme situation (see [41]). Regarding those few data and for the sake of simplicity of the model, we have decided to model people who have suffered this situation by using the specific compartment I_{D_u} . We note that the results reported later in Section 4 indicate that this choice is adequate.

- Quarantine (denoted by Q): The person is in quarantine at home or some special facilities (different from hospitals) during a fixed period.

For example, according to [9, 29], the isolation is performed in one room, always wearing mask at home, etc. to reduce the risk of transmission to cohabiting people. Of course, infection could occur between people in the same home, but this is usually before entering the quarantine period (i.e., while being in the compartment I) and the model takes that into account. Furthermore, according to [19], transmission between people in quarantine at home (thus, aware of their infection) and other person (including cohabiting persons) is severely limited.

Therefore, in this work we assume that a person in compartment Q cannot infect other people (although other possibilities are also easy to include, as explained below) and at the end of this state passes to the compartment R_d . The use of this compartment in the model depends on the strategy of each territory (some of them do not use this control strategy).

- Hospitalized that will recover (denoted by H_R): The person is in hospital and (according to [19, 51, 58]) can still infect other people (in particular healthcare workers). At the end of this state, a person passes to the compartment Q (if implemented in the territory under study; otherwise the person passes directly to the compartment R_d).
- Hospitalized that will die (denoted by H_D): The person is hospitalized and can still infect other people. At the end of this state, a person passes to the compartment D .
- Dead by COVID-19 (denoted by D) after being at hospital.
- Dead by COVID-19 but undetected (denoted by D_u): The person has not survived the disease and the death was not classified by authorities as caused by COVID-19.
- Recovered after being previously detected as infectious (denoted by R_d): The person was previously detected as infectious, survived the disease, passed the quarantine period (if implemented in the territory under study), is no longer infectious and has developed a natural immunity to the virus. Although some cases of reinfection are known, following [47], they seem to be isolated cases; thus, despite the fact that the evolution of this immunity is unknown for a longer period of time, we make the natural immunity assumption. However, it can be modified if new findings regarding this topic are found.
- Recovered after being previously infectious but undetected (denoted by R_u): The person was not previously detected as infectious, survived the disease, is no longer infectious and has developed a natural immunity to the virus (with similar assumptions as above).

We remark that, in this work, we work with two compartments, namely H_R and H_D described previously, to model the dynamics of the number of people in hospital. One may note that a more detailed classification of hospitalization could be considered (for instance, according to the severity of the symptoms, the availability of medical material or beds, etc.) to give more recommendations to decision makers. This could be done by introducing additional compartments to the model. However, this would require the use of detailed data regarding people in hospital which, to the best of our knowledge, are not publicly available for most of the countries (in particular, Italy). Furthermore, adding more compartments will complexify deeply the model, making difficult the parameter estimation process. For those reasons, we decided to use two compartments. We highlight the fact that the results reported in Section 4 (see Figure 2) indicate that this choice is reasonable as the model estimates quite well the reported data related to people in hospital (which are, for the case of Italy, the number of people in hospital, without any specific classification).

The authorities may apply various control measures in order to control the COVID-19 spread (see [6]):

- Isolation: Infected people are isolated from contact with other people. Only health professionals are in contact with them. However, infection of those professionals also occurs (see [64]). Isolated patients receive an adequate medical treatment that reduces the COVID-19 fatality rate.
- Lockdown areas: Movement of people in areas with a high number of infected people may be restricted and controlled, to avoid the spread of the disease.
- Contact tracing: The objective of tracing is to identify potential infectious contacts which may have infected a person or spread COVID-19 to other people. Increase the number of tests in order to increase the percentage of detected infected people.
- Increase of health system resources: The number of operational beds and personal available to detect and treat affected people is increased, producing a decrease in the length of the infectious period for the compartment I .

The model is used to evaluate the spread of a human disease within a considered territory during a fixed time interval.

At the beginning of the simulation, the model parameters are set by the user (as shown below). We can start our simulation at any initial time $t_0 \geq 0$, considering that only susceptible people live in the territories that are free of the disease, whereas the number of people in compartments S , E , I , I_u, I_{D_u} , H_R , H_D , R_d , R_u , D and D_u of the infected territories are set to their corresponding values. Then, for the time interval $[t_0, t_0 + T_{\max}]$, with $T_{\max} \in \mathbb{N}$ being the maximum number of simulation days, the model is applied. If at the end of a simulation day t there is less than 1 person in each compartment E , I , I_u , I_{D_u} , H_R and H_D in all the territory, the simulation is stopped. Else, the simulation is stopped when $t = t_0 + T_{\max}$. Furthermore, the control measures are also implemented and they can be activated or deactivated, when starting the model, in order to quantify their effectiveness to reduce the magnitude and duration of the COVID-19 pandemic in the territory under study.

A diagram summarizing the main structure of our full model is presented in Figure 1.

We assume that people in a territory are characterized to be in one of the compartments described above: S , E , I , I_u , I_{D_u} , H_R , H_D , Q , R_d , R_u , D_u or D . For the sake of simplicity we assume that, at each time, the population inside a territory is homogeneously distributed. Thus, the spatial distribution of the epidemic inside a territory is omitted (it can be taken into account by dividing some territories into a set of smaller regions with similar characteristics). Since the natural natality and mortality (not from COVID-19) do not seem to be important factors for COVID-19 (at least for relatively short periods of time), they are omitted in the model (actually, they can be readily included, if necessary).

In this work we will only consider the within-country/territory disease spread of one territory where, starting from suitable values of t_0 , the COVID-19 pandemic is already spreading by its own, with a relatively negligible dependence on the movement of people in/out of this territory. The between-country/territory disease spread may be also taken into account in the model following the approach developed in [35] with the Be-CoDiS model.

There are some studies about the relevance of humidity and temperature for the spread of COVID-19. In [40], it is shown that the observed patterns of COVID-19 are not completely consistent with the hypothesis that high absolute humidity may limit the survival and transmission of this virus. Furthermore, in [56] it has been found that the lower is the temperature, the greater is the survival period of the SARS-CoV-2 outside the host. Since there is no clear scientific evidence of the effect of the humidity and the temperature on the SARS-CoV-2, we have not included these two factors in our model (this would need to be revised in case of appearing new findings regarding this topic).

Under those assumptions, the evolution of the compartments mentioned above is modeled by the following system

of ordinary differential equations:

$$\begin{aligned}
\dot{S} &= -\frac{S}{N} \left(\beta_E E + \beta_I I + \beta_{I_u} I_u + \beta_{I_{D_u}} I_{D_u} + \beta_{H_R} H_R + \beta_{H_D} H_D \right), \\
\dot{E} &= \frac{S}{N} \left(\beta_E E + \beta_I I + \beta_{I_u} I_u + \beta_{I_{D_u}} I_{D_u} + \beta_{H_R} H_R + \beta_{H_D} H_D \right) - \gamma_E E + \tau_1 - \tau_2, \\
\dot{I} &= \gamma_E E - \gamma_I I, \\
\dot{I}_u &= (1 - \theta - \omega_u) \gamma_I I - \gamma_{I_u} I_u, \\
\dot{I}_{D_u} &= \omega_u \gamma_I I - \gamma_{I_{D_u}} I_{D_u}, \\
\dot{H}_R &= p(\theta - \omega) \gamma_I I - \gamma_{H_R} H_R, \\
\dot{H}_D &= \omega \gamma_I I - \gamma_{H_D} H_D, \\
\dot{Q} &= (1 - p)(\theta - \omega) \gamma_I I + \gamma_{H_R} H_R - \gamma_Q Q, \\
\dot{R}_d &= \gamma_Q Q, \\
\dot{R}_u &= \gamma_{I_u} I_u, \\
\dot{D}_u &= \gamma_{I_{D_u}} I_{D_u}. \\
\dot{D} &= \gamma_{H_D} H_D.
\end{aligned} \tag{1}$$

In this system we have used, as stated above, that people in compartment Q cannot infect other people. To include the possibility of infection, we would only need to add the term $\beta_Q Q$ inside the parenthesis of the first two equations.

In (1):

- N is the number of people in the territory before the start of the pandemic.
- $\omega(t) \in [0, 1]$ is the instantaneous infection detected fatality ratio (iIdFR) in the territory, at time t : the proportion of new detected people (after being in compartment I) that will die of COVID-19 after $\frac{1}{\gamma_{H_D}(t)}$ days, per unit time, compared to the total number of new infectious people (detected or undetected), per unit time. More precisely,

$$\omega(t) = \frac{\dot{D}(t + \frac{1}{\gamma_{H_D}(t)})}{\dot{c}_m(t) + \dot{c}_{m,u}(t)} = \frac{\gamma_{H_D}(t + \frac{1}{\gamma_{H_D}(t)}) H_D(t + \frac{1}{\gamma_{H_D}(t)})}{\gamma_I(t) I(t)},$$

where $c_m(t)$ (respectively, $c_{m,u}(t)$) is the model cumulative number of COVID-19 detected (respectively, undetected) cases at day t (see more details about them below, when describing the outputs of the model).

- $\omega_u(t) \in [0, 1]$ is the instantaneous infection undetected fatality ratio (iIuFR) in the territory, at time t : the proportion of new undetected people (after being in compartment I) that will die of COVID-19 after $\frac{1}{\gamma_{I_{D_u}}(t)}$ days, per unit time, compared to the total number of new infectious people (detected or undetected), per unit time. More precisely,

$$\omega_u(t) = \frac{\dot{D}_u(t + \frac{1}{\gamma_{I_{D_u}}(t)})}{\dot{c}_m(t) + \dot{c}_{m,u}(t)}.$$

We point out that $\omega(t) + \omega_u(t) \in [0, 1]$ is the instantaneous infection fatality ratio (iIFR), at time t : the proportion of new (detected or undetected) people (after being in compartment I) that will die of COVID-19, per unit time, compared to the total number of new infectious people (detected or undetected), per unit time. This value is closely related to the IFR defined below (see the **Results** section). Actually, if it is assumed that the number of total deaths is proportional to the number of total cases, then IFR is, essentially, $\omega + \omega_u$.

- $\theta(t) \in [\omega(t), 1 - \omega_u(t)] \subset [0, 1]$ is the proportion of the number of new infectious people that are detected and documented by the authorities (after being in compartment I), per unit time, compared to the total number of new infectious people (detected or undetected), per unit time, at time t . More precisely,

$$\theta(t) = \frac{\dot{c}_m(t)}{\dot{c}_m(t) + \dot{c}_{m,u}(t)}.$$

We point out that

$$\frac{\omega(t)}{\theta(t)} = \frac{\dot{D}(t + \frac{1}{\gamma_{H_D}(t)})}{\dot{c}_m(t)} \in [0, 1]$$

is the instantaneous case fatality ratio (iCFR), at time t : the proportion of the number of new detected people (after being in compartment I) that will die of COVID-19 after $\frac{1}{\gamma_{H_D}(t)}$ days, per unit time, compared to the number of new detected people (after being in compartment I), per unit time.

This value is closely related to the Case Fatality Ratio (CFR) defined as the expected value of the probability of dying for a person who is infected and detected, i.e. the number of detected deaths due to the virus over the number of detected infections. Actually, if it is assumed that the number of detected deaths is proportional to the number of detected cases, then CFR is, essentially, $\frac{\omega}{\theta}$.

- $\beta_E, \beta_I, \beta_Q, \beta_{I_u}, \beta_{I_{D_u}}, \beta_{H_R}, \beta_{H_D} \in [0, \infty)$ are the disease contact rates (day^{-1}) of a person in the respective compartments. They may change with time, due to the control measures.
- $\gamma_I(t) \in (0, +\infty)$ is the transition rate (day^{-1}) from compartment I to compartments I_u, I_{D_u}, Q, H_R or H_D , at time t . Similarly, $\gamma_E, \gamma_{I_u}(t), \gamma_{I_{D_u}}(t), \gamma_Q(t), \gamma_{H_R}(t)$, and $\gamma_{H_D}(t) \in (0, +\infty)$ denote the transition rates (day^{-1}) from compartments E, I_u, I_{D_u}, Q, H_R and H_D to compartments I, R_u, D_u, R_d, Q and D , respectively, at time t . For some particular territories, some compartments could be excluded from the model.
- p is the ratio of the number of new detected infected people that will survive the disease and are hospitalized (they go to compartment H_R) over the total number of new detected infected people that will survive the disease (and may also do quarantine, compartment Q , without going to hospital).
- $\tau_1(t)$ and $\tau_2(t)$ are the people infected that arrive/leave from/to other territories, per day. Both can be modeled following the between-country spread part of the Be-CoDiS model described in [35]. For the sake of simplicity we have included those terms only in the second equation, corresponding to compartment E , but they could be included in the equations for other compartments too. When a discrete sequence of imported/exported infections is known it can be included in the numerical resolution of the system by simple addition/removal to/from the suitable compartment, at the corresponding discrete time.

System (1) is completed with initial data $S(t_0), E(t_0), I(t_0), I_u(t_0), I_{D_u}(t_0), Q(t_0), H_R(t_0), H_D(t_0), R_d(t_0), R_u(t_0), D_u(t_0)$ and $D(t_0)$ given in $[0, \infty)$. These values must be non-negative and $S(t_0) + E(t_0) + I(t_0) + I_u(t_0) + I_{D_u}(t_0) + Q(t_0) + H_R(t_0) + H_D(t_0) + R_d(t_0) + R_u(t_0) + D_u(t_0) + D(t_0) = N$. It is standard (see, e.g., [34], [36]) to prove that $S(t), E(t), I(t), I_u(t), I_{D_u}(t), Q(t), H_R(t), H_D(t), R_d(t), R_u(t), D_u(t)$ and $D(t)$ are non-negative and $S(t) + E(t) + I(t) + I_u(t) + I_{D_u}(t) + Q(t) + H_R(t) + H_D(t) + R_d(t) + R_u(t) + D_u(t) + D(t) = N$, for all $t \geq t_0$.

We remark that system (1) is an epidemic-type model (as described, e.g., in [22]), which is not trying to capture endemic behaviors. Actually, there is no incoming flux in the Susceptible compartment; therefore, the fate of every individual will be either staying susceptible, recovering or dying. This means there are no endemic equilibria in our system; on the contrary, there exists a manifold $S + R_d + R_u + D_u + D = N$ containing infinite disease-free equilibrium points.

We point out that the 8th, 9th, 10th, 11th and 12th equations of system (1) are not coupled with the other equations. Thus, we can solve the first seven equations of that system and the solution of the last five ones can be computed as follows: Q is the solution of the 8th equation of system (1), which is a linear ordinary differential equation (we do not include here the easy corresponding explicit solution, for the sake of simplicity) and

$$\begin{aligned} D(t) &= D(t_0) + \int_{t_0}^t \gamma_{H_D}(s)H_D(s)ds, \\ D_u(t) &= D_u(t_0) + \int_{t_0}^t \gamma_{I_{D_u}}(s)I_{D_u}(s)ds, \\ R_d(t) &= R_d(t_0) + \int_{t_0}^t \gamma_Q(s)Q(s)ds, \\ R_u(t) &= R_u(t_0) + \int_{t_0}^t \gamma_{I_u}(s)I_u(s)ds. \end{aligned}$$

It is important to remark that, if for some territory the model does not include the compartment Q , then $p = 1$, $\dot{R}_d = \gamma_{H_R}H_R$ and

$$R_d(t) = R_d(t_0) + \int_{t_0}^t \gamma_{H_R}(s)H_R(s)ds.$$

For the numerical simulations presented in this paper, the first eight equations of system (1) were numerically solved with the classic 4th order Runge–Kutta method with four stages (RK4) and with four hours as time step (which was tested to see that it is suitable to get stable results).

2.2 Outputs of the model

Here, we present some of the outputs that can be obtained from the mathematical model, used to analyze the results of the simulations:

- $c_m(t)$: The model cumulative number of COVID-19 detected cases at day t , which is given by

$$c_m(t) = H_R(t) + H_D(t) + Q(t) + R_d(t) + D(t),$$

and can be also computed as follows:

$$c_m(t) = c_m(t_0) + \int_{t_0}^t \frac{d}{dt} \left(H_R(s) + H_D(s) + Q(s) + R_d(s) + D(s) \right) ds = c_m(t_0) + \int_{t_0}^t \theta(s)\gamma_I(s)I(s)ds.$$

- $c_{m,n}(t)$: The number of new cases per day at day t , which is given by

$$c_{m,n}(t) = \theta(t)\gamma_I(t)I(t).$$

- $h_m(t)$: The model cumulative number of COVID-19 detected infected healthcare workers at day t , which is given by

$$h_m(t) = h_m(t_0) + \int_{t_0}^t \frac{S(s)}{N} \left(\beta_{H_R}(s)H_R(s) + \beta_{H_D}(s)H_D(s) \right) ds.$$

- $c_{m,u}(t)$: The model cumulative number of COVID-19 undetected cases at day t , which is given by

$$c_{m,u}(t) = E(t) + I(t) + I_u(t) + I_{D_u}(t) + R_u(t) + D_u(t),$$

and can be also computed as follows:

$$\begin{aligned}
c_{m,u}(t) &= E(t) + I(t) + I_u(t_0) + I_{D_u}(t_0) + R_u(t_0) + D_u(t_0) + \int_{t_0}^t \frac{d}{dt} \left(I_u(s) + I_{D_u}(s) + R_u(s) + D_u(s) \right) ds \\
&= E(t) + I(t) + I_u(t_0) + I_{D_u}(t_0) + R_u(t_0) + D_u(t_0) + \int_{t_0}^t (1 - \theta(s)) \gamma_I(s) I(s) ds.
\end{aligned}$$

- $d_m(t)$: The model cumulative number of detected deaths (due to COVID-19), at day t , which is given by $D(t)$.
- $d_{m,u}(t)$: The model cumulative number of undetected deaths, at day t , which is given by $D_u(t)$.
- R_e : The effective reproduction number of COVID-19. It is defined as the number of cases that one infected person generates, on average, over the course of its infectious period [5]. Part of the population can be already infected and/or control measures may have been implemented. It depends on the considered territory and changes during the spread of the disease. Furthermore, if at time $t=0$ no control measures have been applied, $R_e(0) = R_0$, where R_0 is the basic reproduction number of COVID-19, which is defined as the number of cases that one infected person generates on average over the course of its infectious period, in an otherwise uninfected population and without special control measures. Typically, the spread of the disease slows down when $R_e(t) < 1$. Taking into account this definition, for cases with constant transition rates we propose the following non-classical estimation:

$$\begin{aligned}
R_e(t) &= \int_t^{t+\frac{1}{\gamma_E}} \beta_E(\tau) \frac{S(\tau)}{N} d\tau + \int_{t+\frac{1}{\gamma_E}}^{t+\frac{1}{\gamma_E}+\frac{1}{\gamma_I}} \beta_I(\tau) \frac{S(\tau)}{N} d\tau \\
&+ \left(1 - \theta\left(t + \frac{1}{\gamma_E} + \frac{1}{\gamma_I}\right) - \omega_u\left(t + \frac{1}{\gamma_E} + \frac{1}{\gamma_I}\right) \right) \int_{t+\frac{1}{\gamma_E}+\frac{1}{\gamma_I}}^{t+\frac{1}{\gamma_E}+\frac{1}{\gamma_I}+\frac{1}{\gamma_{I_u}}} \beta_{I_u}(\tau) \frac{S(\tau)}{N} d\tau \\
&+ \omega_u\left(t + \frac{1}{\gamma_E} + \frac{1}{\gamma_I}\right) \int_{t+\frac{1}{\gamma_E}+\frac{1}{\gamma_I}}^{t+\frac{1}{\gamma_E}+\frac{1}{\gamma_I}+\frac{1}{\gamma_{I_{D_u}}}} \beta_{I_{D_u}}(\tau) \frac{S(\tau)}{N} d\tau \\
&+ p\left(t + \frac{1}{\gamma_E} + \frac{1}{\gamma_I}\right) \left(\theta\left(t + \frac{1}{\gamma_E} + \frac{1}{\gamma_I}\right) - \omega\left(t + \frac{1}{\gamma_E} + \frac{1}{\gamma_I}\right) \right) \int_{t+\frac{1}{\gamma_E}+\frac{1}{\gamma_I}}^{t+\frac{1}{\gamma_E}+\frac{1}{\gamma_I}+\frac{1}{\gamma_{H_R}}} \beta_{H_R}(\tau) \frac{S(\tau)}{N} d\tau \\
&+ \omega\left(t + \frac{1}{\gamma_E} + \frac{1}{\gamma_I}\right) \int_{t+\frac{1}{\gamma_E}+\frac{1}{\gamma_I}}^{t+\frac{1}{\gamma_E}+\frac{1}{\gamma_I}+\frac{1}{\gamma_{H_D}}} \beta_{H_D}(\tau) \frac{S(\tau)}{N} d\tau.
\end{aligned} \tag{2}$$

In this formula, we estimate the mean number of secondary cases produced by one person infected at time t , during its whole infectiousness period, taking into account all the possible evolutions of its health status.

Composite rectangular, trapezoidal or Simpson rules can be used to obtain an estimation of (2).

We remark that other estimations for R_0 can be also considered. For instance, one can use the so-called "Next generation Matrix" method to approximate its value (see [54]). However, as pointed in [7], this approximation does not correspond accurately to the biological definition of R_0 (e.g., parameters are assumed to be constant and the change of the proportion of susceptible people during the infectious time of a person is omitted). We remark that R_e should be used with caution by policymakers, without relying only on that value (see [1]).

Note that we can estimate the particular contribution of each infectious compartment (i.e. E , I , I_u , I_{D_u} , H_R and H_D) to the value of R_e and R_0 by computing only their respective integrals in (2).

- $Hos(t)$: The number of people in hospital is estimated as follows:

$$Hos(t) = H_D(t) + H_R(t).$$

This function can help to estimate and plan the number of clinical beds needed to treat all the COVID-19 cases.

- $Hos_n(t)$: The number of new hospitalized people per day at day t , which is given by

$$Hos_n(t) = [p(\theta(t) - \omega(t)) + \omega(t)] \gamma_I(t) I(t).$$

- MHos: The maximum number of hospitalized people at the same time (in territory i) during the time interval $[t_0, T]$. It is computed as:

$$\text{MHos} = \max_{t \in [t_0, T]} \text{Hos}(t).$$

This number can help to estimate and plan the number of clinical beds needed to treat all the COVID-19 cases.

- $\Gamma_E(t)$ and $\Gamma_{I_u}(t)$: The number of people infected during the time interval $[t_0, T]$, by contact with people in compartment E and I_u , respectively. They are given by:

$$\begin{aligned} \Gamma_E(t) &= \int_{t_0}^t \beta_E(s) E(s) \frac{S(s)}{N} ds, \\ \Gamma_{I_u}(t) &= \int_{t_0}^t \beta_{I_u}(s) I_u(s) \frac{S(s)}{N} ds. \end{aligned}$$

We point out that those outputs and others that can be computed similarly can be compared with reported data when available.

3 Model parameter estimation for COVID-19

Here, we present the methodology we used to estimate the parameters of the model described in Section 2.1. We detail each kind of parameter according to its category.

3.1 Parametrization of some of the functions involved in the model

The model involves several functions. We now detail how some of these functions can be written, using empirical assumptions, in terms of constant parameters, which will have to be later identified for each territory under study (we will do it below for the case of Italy). Of course, other alternatives may be also used, taking into account the particularities of each territory.

-Disease contact rates: We consider that

$$\begin{aligned} \beta_E(t) &= m_E(t) \beta_{E,0}, \quad \beta_I(t) = m_I(t) \beta_{I,0}, \quad \beta_{I_u}(t) = m_{I_u}(t) \beta_{I_u,0}(\theta(t)), \quad \beta_{I_{D_u}}(t) = m_{I_{D_u}}(t) \beta_{I_{D_u},0}, \\ \beta_{H_R}(t) &= m_{H_R}(t) \beta_{H_R,0}, \quad \beta_{H_D}(t) = m_{H_D}(t) \beta_{H_D,0}. \end{aligned}$$

Here $\beta_{E,0}, \beta_{I,0}, \beta_{I_u,0}, \beta_{I_{D_u},0}, \beta_{H_R,0}, \beta_{H_D,0}$ are the respective disease contact rates when no control measures are taken and $m_E, m_I, m_{I_u}, m_{I_{D_u}}, m_{H_R}, m_{H_D}$ are functions simulating de control measures, as described below.

Regarding $\beta_{I_u,0}$, we may consider

$$\beta_{I_u,0}(\theta) = \begin{cases} \bar{\beta}_{I,0} & , \text{ if } \theta \in [0, \bar{\omega}], \\ \text{nonincreasing} & , \text{ if } \theta \in (\bar{\omega}, 1), \\ \underline{\beta}_{I,0} & , \text{ if } \theta = 1, \end{cases}$$

where $\bar{\omega}$ is a fatality rate value that is explained below and $\underline{\beta}_I$ and $\bar{\beta}_I$ are suitable lower and upper bounds, respectively. For the sake of simplicity, here we take $\bar{\beta}_{I,0} = \beta_{I,0}$.

Furthermore, following the idea proposed in [33, 35, 39], we assume that there exists a relationship between the contact rates $\beta_{I,0}, \beta_{E,0}, \beta_{I_{D_u},0}$ and $\underline{\beta}_{I,0}$.

More precisely, following [33], we consider that people in compartments E, I_u and I_{D_u} are less infectious than people in compartment I (due to their probably lower virus load or isolation measures, see [2, 39, 61]). Thus, we consider that $\beta_{E,0} = C_E \beta_{I,0}$ and $\underline{\beta}_{I,0} = C_u \beta_{I,0}$ and $\beta_{I_{D_u},0} = C_{I_{D_u},0} \beta_{I,0}$, where $C_E, C_u, C_{I_{D_u},0} \in [0, 1]$.

A similar approach can be followed for β_{H_R} and β_{H_D} assuming, for instance, that $\beta_{H_R,0} = \beta_{H_D,0} = C_H \beta_{I,0}$. Alternatively (as we do here for the case of Italy) we can compute good estimates for β_{H_R} and $\beta_{H_D,0}$ using historical data, as shown below.

-Control measures: We may distinguish three different types of control measures: Social distancing, contact tracing and health system measures. Each of them may be included in the model by using different functions, as explained below.

Functions $m_E(t)$, $m_I(t)$, $m_{I_u}(t) \in [0, 1]$ (%) represent the reduction of some of the disease contact rates, at time t , due to the efficiency of the control measures corresponding to social distancing applied to the compartments E , I and I_u , respectively (see, e.g., [46]). Here, for the sake of simplicity, we consider that those three functions are equal (of course, other alternatives are also possible) and given by (see [33])

$$m_E(t) = m_{I_u}(t) = m_I(t) = m(t) = \begin{cases} (m_0 - m_1) \exp\left(-\kappa_1(t - t_0)\right) + m_1 & t \in [t_0, \lambda_1], \\ (m(\lambda_1) - m_2) \exp\left(-\kappa_2(t - \lambda_1)\right) + m_2 & t \in (\lambda_1, \lambda_2], \\ \vdots \\ (m(\lambda_{q-1}) - m_q) \exp\left(-\kappa_q(t - \lambda_{q-1})\right) + m_q & t \in (\lambda_{q-1}, \infty). \end{cases}$$

We may also define in a similar way functions $m^{(t)}$ and $m^{(s)}$ related with contact tracing and health system control measures, respectively, which include similar parameters denoted by $m_i^{(t)}$, $m_i^{(s)}$, $\kappa_i^{(t)}$, $\kappa_i^{(s)}$, $q^{(t)}$ and $q^{(s)}$. Here, $q, q^{(t)}, q^{(s)} \in \mathbb{N}$ are the number of changes of control strategies.

Here, for every $i \in \{0, \dots, q\}$, $j \in \{0, \dots, q^{(s)}\}$ and $k \in \{0, \dots, q^{(t)}\}$, constants $m_i, m_j^{(s)}, m_k^{(t)} \in [0, 1]$ measure the intensity of the control measures (lower value implies higher intensity), $\kappa_i, \kappa_j^{(s)}, \kappa_k^{(t)} \in [0, +\infty)$ (day⁻¹) simulate the efficiency of the control measures (higher value implies faster effectiveness of the control measures) and $\lambda_i, \lambda_j^{(s)}, \lambda_k^{(t)} \in [t_0, \infty]$ denote days of changes of each control strategy.

The values of $\lambda_i, \lambda_j^{(s)}$ are typically taken from the literature (using the dates when the territories implement special control measures). Some of the values of $m_i, m_j^{(s)}$ can be also known. The rest of the parameters need to be identified, as explained below. When doing this, to ease the search process, we use new parameters c_i instead of m_i , as follows:

- If we want $m_{i+1} \leq m_i$, then we look for a suitable parameter $c_{i+1} \in [0, 1]$ such that $m_{i+1} = c_{i+1} m_i$.
- If we want $m_{i+1} \geq m_i$, then we look for a suitable parameter $c_{i+1} \in [0, 1]$ such that $m_{i+1} = m_i + c_{i+1}(1 - m_i)$.

A similar approach can be followed by using new parameters $c_i^{(s)}$ instead of $m_i^{(s)}$.

Function θ is related with the contact tracing measures, since the higher those measures are at time t , the closer $\theta(t)$ is to 1. The other function involved with the contact tracing measures is $m^{(t)}$, which is used below for estimating the transition rates. The function $m^{(s)}$ is used to compute the fatality ratios and to define $m_{H_R}(t) = m_{H_D}(t) = m^{(s)}(t) \in [0, 1]$, representing the reduction of the disease contact rates $\beta_{H_R,0}, \beta_{H_D,0}$, at time t , due to the efficiency of the health system measures applied to the corresponding compartments. We note that this approach is not taken into account when using the alternative mentioned above to compute estimates for β_{H_R} and β_{H_D} using historical data (which will be shown below).

-Fatality ratios: As observed in other epidemics (see, e.g. [35]), $\omega(t)$ can be affected by the application of health system measures. Thus, following [33], we propose to consider

$$\omega(t) = m^{(s)}(t)\bar{\omega} + (1 - m^{(s)}(t))\underline{\omega},$$

where $\bar{\omega} \in [0, 1]$ is the iIdFR when no health system measures are applied (i.e. $m^{(s)} = 1$); and $\underline{\omega} \in [0, 1]$ is the iIdFR when the implemented health system measures are fully applied (i.e. $m^{(s)} = 0$). We note that $\bar{\omega} \geq \underline{\omega}$, thus, we consider

$\bar{\omega} = \underline{\omega} + \delta_\omega$, with $\delta_\omega \in [0, 1 - \underline{\omega}]$. Alternatively (as we do here for the case of Italy), these constants can be estimated from values of the IFR in the literature (see, e.g., [3, 53]).

Additionally, we assume that

$$\omega_u(t) = \begin{cases} \omega_{u,0} & , \text{ if } t = t_0 \\ \omega_{u,1} & , \text{ if } t = t_{u,1} \\ \vdots & \\ \omega_{u,n} & , \text{ if } t = t_{u,n} \\ \text{constant after the last value } t_{u,n} & \\ \text{linear otherwise,} & \end{cases}$$

where $n \in \mathbf{N} \cup \{0\}$ and the corresponding values t_i , $\omega_{u,i}$ are parameters to be chosen for each territory, depending on its particular circumstances.

-Transition rates:

- According to [39], the transition rate γ_E from E to I depends only on the disease and, therefore, is considered constant.
- Some other transition rates could depend on time. For instance, according to [39, 38], depending on the studied territory, the value of $\gamma_I(t)$ can be increased due to the application of contact tracing control measures (i.e. people with symptoms are detected earlier). As a consequence, the values of $\gamma_{I_u}(t)$, $\gamma_{I_{D_u}}(t)$, $\gamma_Q(t)$, $\gamma_{H_R}(t)$ and $\gamma_{H_D}(t)$ can be decreased (e.g. people with symptoms stay under observation during more time). In order to parametrize these functions, we denote by d_I , d_{I_u} , $d_{I_{D_u}}$, d_Q , d_{H_R} and d_{H_D} the initial duration (given as a suitable average) in days of a person in compartment I , I_u , I_{D_u} , Q , H_R and H_D , respectively, without the application of special contact tracing measures. Then, following [35], we may use the following functions:

$$\begin{aligned} \gamma_I(t) &= \frac{1}{d_I - g(t)} \text{ (day}^{-1}\text{)}, & \gamma_{I_u}(t) &= \frac{1}{d_{I_u} + g(t)} \text{ (day}^{-1}\text{)}, & \gamma_{I_{D_u}}(t) &= \frac{1}{d_{I_{D_u}} + g(t)} \text{ (day}^{-1}\text{)}, \\ \gamma_Q(t) &= \frac{1}{d_Q + g(t)} \text{ (day}^{-1}\text{)}, & \gamma_{H_R}(t) &= \frac{1}{d_{H_R} + g(t)} \text{ (day}^{-1}\text{)}, & \gamma_{H_D}(t) &= \frac{1}{d_{H_D} + g(t)} \text{ (day}^{-1}\text{)}, \end{aligned}$$

where $g(t) = d_g(1 - m^{(t)}(t))$ represents the decrease of the duration of d_I due to the application of contact tracing measures, at time t ; and d_g is the maximum number of days that d_I can be decreased.

We remark that, for some territories, it can be suitable to consider constant (i.e. not dependent on time) all the transition rates, which corresponds to the case with $d_g = 0$.

-Ratio p : We describe it by using the following function depending on the value of $\theta - \omega$ (we remark that other alternatives are also possible):

$$p = \begin{cases} \frac{p_0(\theta_0 - \omega_0)}{\theta - \omega} & , \text{ if } \theta - \omega \geq \theta_0 - \omega_0, \\ 1 - \frac{1 - p_0}{\theta_0 - \omega_0}(\theta - \omega) & , \text{ if } \theta - \omega \leq \theta_0 - \omega_0. \end{cases}$$

Here $p_0 = p(t_{\theta_0})$ and $\omega_0 = \omega(t_{\theta_0})$, where t_{θ_0} and θ_0 are described below (when we show how to estimate $\theta(t)$). Indeed, we assume that, if $\theta - \omega > \theta_0 - \omega_0$ all new detected infected people, when comparing to the case of θ_0 and ω_0 , will go to state Q , as they should present light symptoms. Furthermore, we use a linear interpolation between the case $\theta - \omega = 0$ (with $p = 1$, i.e., only persons with severe symptoms are detected) and $\theta - \omega = \theta_0 - \omega_0$ (with $p = p_0$). More complex formulas could be also considered, but we use the previous one, for the sake of simplicity.

We remark that one of the functions involved in the model, namely $\theta(t)$, has not been parameterized above. We will show below how to deal with it.

3.2 Estimation of parameters and non-parameterized functions involved in the model

Once the model is built, if we want to use it for a particular country or territory, we need to identify suitable model parameters and functions for the case under study. Some of them can be found in the literature. However, due to lack of information regarding the behavior of the SARS-CoV-2 and implicit uncertainty, others need to be estimated.

To do that, we need as much information as possible about the dynamics of the pandemic in the studied territory. In particular, we may use the following temporal series of the territory under study, if available:

- $c_r(t)$: The reported cumulative number of COVID-19 infected, cases at time t .
- $d_r(t)$: The reported cumulative number of COVID-19 deaths, at time t .
- $h_r(t)$: The reported cumulative number of COVID-19 infected healthcare workers, at time t .
- $\tau_{1,r}(t)$: The reported number of known imported cases on day t . More precisely, infected people arriving from an external territory. Usually they are in the incubation period, but, in some cases, they are moving between hospitals of different countries or territories (this last possibility is not included in the equations, for the sake of simplicity, but it can be easily done).
- $\tau_{2,r}(t)$: The reported number of known exported cases on day t . More precisely, infected people going to an external territory. Usually, they are in the incubation period, but sometimes they are moving between hospitals of different countries or territories (this last possibility is not included in the equations, for the sake of simplicity, but it can be easily done).
- $d_{r,u}(t)$: The reported cumulative number of estimated undetected COVID-19 deaths, at time t .
- $Hos_r(t)$: The reported number of people in hospital (i.e. Hos_r), at time t .
- $q_r(t)$: The number of people in quarantine at home, due to their infection with COVID-19.

We note that, sometimes, there is only partial information and the gaps can be filled with suitable techniques (see, e.g., [17]).

Now, with that information we may estimate some of the model parameters and functions, as follows:

-Ratio $\theta(t)$ of new infected people that are detected: To estimate $\theta(t)$, first, we compute an estimation of iCFR ($\frac{\omega(t)}{\theta(t)}$) as follows:

$$\frac{\omega(t)}{\theta(t)} = \begin{cases} \frac{d_r(t_{iCFR} + \frac{1}{\gamma_{HD}})}{c_r(t_{iCFR})} & , \text{ if } t \leq t_{iCFR}, \\ \frac{d_r(t + \frac{1}{\gamma_{HD}}) - d_r(t - r + \frac{1}{\gamma_{HD}})}{c_r(t) - c_r(t - r)} & , \text{ if } t_{iCFR} < t < T_{\max} - \frac{1}{\gamma_{HD}}, r \text{ is the minimum natural number such that } \\ & c_r(t) - c_r(t - r) \neq 0 \text{ and this quotient is } \leq 1, \\ 1 & , \text{ if } t_{iCFR} < t < T_{\max} - \frac{1}{\gamma_{HD}}, r \text{ is the minimum natural number such that } \\ & c_r(t) - c_r(t - r) \neq 0 \text{ and the previous quotient is } > 1, \\ \text{linear interpolation} & , \text{ otherwise,} \end{cases}$$

where $t_{iCFR} \geq 7$ days is the first date for which $c_r(t) - c_r(t - 1) \neq 0$ and $d_r(t + \frac{1}{\gamma_{HD}}) \neq 0$. We note that, in this formula, we estimate the iCFR by taking into account the delay between the detection of a case and its death (instead of the ratio between cases and deaths reported at the same date). This delay could also be considered in the system of differential equations, instead of the exponential-decay-formulation, leading to a system of delay differential equations (see, e.g., [23]).

After that, to regularize the obtained iCFR curve, we apply a filter that consists in computing the average value of the last week. We denote the resulting filtered curve by $\omega_{\text{CFR}}(t)$:

$$\omega_{\text{CFR}}(t) = \begin{cases} \frac{\omega(t_{\text{iCFR}})}{\theta(t_{\text{iCFR}})} & , \text{ if } t \leq t_{\text{iCFR}}, \\ \frac{1}{7} \sum_{i=0}^6 \frac{\omega(t-i)}{\theta(t-i)} & , \text{ if } t > t_{\text{iCFR}}. \end{cases} \quad (3)$$

To avoid unrealistic situations (due to poor data), if $\omega_{\text{CFR}}(t) \leq 0.01$ (which is assumed too low, see, e.g., [53]) for some $t > t_{\text{iCFR}}$, then we extend the average of the last week used to compute this $\omega_{\text{CFR}}(t)$ by including the previous days until getting $\omega_{\text{CFR}}(t) > 0.01$.

Finally, we define $\theta(t)$ by:

$$\theta(t) = \begin{cases} \theta_0 & , \text{ if } t \leq t_{\theta_0}, \\ \frac{\omega(t)}{\omega_{\text{CFR}}(t)} & , \text{ if } t > t_{\theta_0}, \end{cases}$$

where $\theta_0 = \frac{\omega(t_{\theta_0})}{\omega_{\text{CFR}}(t_{\theta_0})}$ and $t_{\theta_0} = t_{\text{iCFR}} + 6$ days to guarantee that $\omega_{\text{CFR}}(t_{\theta_0})$ is computed by using the average value defined by the second line of (3), for the sake of smoothness. Notice that, if we simulate until the current date, we do not have enough data to compute $\frac{\omega(t)}{\theta(t)}$. In this case, we set $\theta(t)$ to a constant value for the last days of the simulation without data. When simulating scenarios for future dates, we can also keep the last value of θ or use alternative possible values.

For particular territories (as we do here for the case of Italy) we can use different values of IFR found in the literature and use them as ω in the previous formula. That will provide different estimations for $\theta(t)$.

-Contact rates $\beta_{H_R}(t)$ and $\beta_{H_D}(t)$: We use the following estimation of the fraction of infected healthcare workers (assumed to be due to contact with hospitalized people, see, e.g., [58]):

$$\frac{\left(p(\theta - \omega) \frac{\beta_{H_R}}{\gamma_{H_R}} + \omega \frac{\beta_{H_D}}{\gamma_{H_D}} \right)}{\left(p(\theta - \omega) \frac{\beta_{H_R}}{\gamma_{H_R}} + \omega \frac{\beta_{H_D}}{\gamma_{H_D}} \right) + \frac{\beta_E}{\gamma_E} + \frac{\beta_I}{\gamma_I} + (1 - \theta - \omega_u) \frac{\beta_{I_u}}{\gamma_{I_u}} + \omega_u \frac{\beta_{I_{D_u}}}{\gamma_{I_{D_u}}} } = \eta.$$

To compute η , we first define $\tilde{\eta}$ as

$$\tilde{\eta}(t) = \begin{cases} \frac{h_r \left(t_\eta + \frac{1}{\gamma_E} + \frac{1}{\gamma_I} \right)}{c_r \left(t_\eta + \frac{1}{\gamma_E} + \frac{1}{\gamma_I} \right)} & , \text{ if } t \leq t_\eta, \\ \frac{h_r \left(t + \frac{1}{\gamma_E} + \frac{1}{\gamma_I} \right) - h_r \left(t - r + \frac{1}{\gamma_E} + \frac{1}{\gamma_I} \right)}{c_r \left(t + \frac{1}{\gamma_E} + \frac{1}{\gamma_I} \right) - c_r \left(t - r + \frac{1}{\gamma_E} + \frac{1}{\gamma_I} \right)} & , \text{ if } t_\eta < t < T_{\max} - \frac{1}{\gamma_E} - \frac{1}{\gamma_I}; r \text{ is the minimum natural number} \\ & \text{such that } h_r \left(t + \frac{1}{\gamma_E} + \frac{1}{\gamma_I} \right) \text{ and } h_r \left(t - r + \frac{1}{\gamma_E} + \frac{1}{\gamma_I} \right) \text{ are known} \\ & \text{and } c_r \left(t + \frac{1}{\gamma_E} + \frac{1}{\gamma_I} \right) - c_r \left(t - r + \frac{1}{\gamma_E} + \frac{1}{\gamma_I} \right) \neq 0; \\ & \text{and this quotient is } \leq 1, \\ 1 & , \text{ under the same conditions as in the previous line but when} \\ & \text{the quotient is } > 1, \\ \text{linear interpolation} & , \text{ otherwise,} \end{cases}$$

and t_η is the first date for which $h_r \left(t_\eta + \frac{1}{\gamma_E} + \frac{1}{\gamma_I} \right) \neq 0$ is reported. We point out that, in the previous formula, we have made a distinction about the knowledge of the values of h_r , because they are rarely given every day. We remark that we are taking into account the delay of $\frac{1}{\gamma_E} + \frac{1}{\gamma_I}$ days between the infection of a person and its detection by authorities. Thus, we set $\tilde{\eta}(t)$ to a constant value for the last $\frac{1}{\gamma_E} + \frac{1}{\gamma_I}$ days. Finally, to avoid unrealistic situations (due to poor data), we take the following seven-day average value:

$$\eta(t) = \begin{cases} \frac{1}{7} \sum_{i=0}^6 \tilde{\eta}(t_0 + i) & , \text{ if } t < t_0 + 3, \\ \frac{1}{7} \sum_{i=-3}^3 \tilde{\eta}(t + i) & , \text{ if } t_0 + 3 \leq t \leq t_0 + T_{\max} - 3, \\ \frac{1}{7} \sum_{i=0}^6 \tilde{\eta}(t_0 + T_{\max} - i) & , \text{ if } t > t_0 + T_{\max} - 3. \end{cases} \quad (4)$$

Assuming, for the sake of simplicity, that $\beta_{H_R} = \beta_{H_D}$, we have that

$$\beta_{H_R} = \beta_{H_D} = \frac{\eta \left(\frac{\beta_E}{\gamma_E} + \frac{\beta_I}{\gamma_I} + (1 - \theta - \omega_u) \frac{\beta_{I_u}}{\gamma_{I_u}} + \omega_u \frac{\beta_{I_{D_u}}}{\gamma_{I_{D_u}}} \right)}{(1 - \eta) \left(p(\theta - \omega) \frac{1}{\gamma_{H_R}} + \omega \frac{1}{\gamma_{H_D}} \right)}.$$

In the previous expression, the coefficients depend on time t .

-Other unknown parameters: We can identify them with a multiobjective technique, as described below.

3.3 Identification of some parameters with a multiobjective optimization technique

First, we determine the set of unknown model parameters to be identified by the approach presented here (which will depend on the available information in the considered territory). For each of those parameters, we establish a suitable interval of possible values (given by upper and lower bounds of each parameter) and denote by $S \subset \mathbb{R}^n$ the feasible region given by the cartesian product of all those intervals, where n is the number of unknown parameters.

Once S is defined, we follow the methodology proposed in [15]. This approach aims at minimizing the differences between the temporal series reported by the authorities and the corresponding ones obtained as model outputs when considering particular values $\Omega \in S$. Then, given two weight coefficients $\rho_1 > 0$ and $\rho_2 > 0$ and a temporal series X , we define the function

$$f_X(\Omega) = \frac{\|X_r - X_m^\Omega\|_{L^2(t_0, t_0 + T_{\max})}^2 + \rho_1 \|\max(X_r - X_m^\Omega, 0)\|_{L^2(t_0, t_0 + T_{\max})}^2}{\|X_r\|_{L^2(t_0, t_0 + T_{\max})}^2} + \rho_2 \frac{|\max(X_r(t_0 + T_{\max}) - X_m^\Omega(t_0 + T_{\max}), 0)|^2}{|X_r(t_0 + T_{\max})|^2}, \quad (5)$$

that measures the error between the reported data, which is denoted by X_r , and the model output denoted by X_m^Ω . Notice that, as done in previous works [15, 33], it is based on the relative error. However, as a novelty, we add two new terms to penalize those Ω obtaining a temporal series that underestimate the reported data. We recall that, to reduce the risk of over-fitting to the reported data [50], we use the L^2 norm in $[t_0, t_0 + T_{\max}]$, which is

$$\|g\|_{L^2([t_0, t_0 + T_{\max}])} = \left(\int_{t_0}^{t_0 + T_{\max}} |g(t)|^2 dt \right)^{1/2}.$$

The number of functions $f_X(\Omega)$ depends on the number of available temporal series of the territory under study. The authorities usually report at least the cumulative daily number of cases and deaths (i.e. c_r and d_r), see, e.g., [62]. In some cases, they also include other data as, for instance, the number of people in hospital (i.e. Hos_r), in quarantine at home (i.e. q_r), or the cumulative number of undetected deaths (i.e. $d_{r,u}$) (see, e.g., [44]).

Then, we formulate a multiobjective problem that may have as many objective functions as available temporal series. For instance, we could solve

$$\begin{cases} \min \{f_c(\Omega), f_d(\Omega), f_{\text{Hos}}(\Omega), f_{d_u}(\Omega)\} \\ \text{subject to } \Omega \in S \subset \mathbb{R}^n, \end{cases} \quad (6)$$

where S is the feasible region (i.e. upper and lower bounds of each parameter) and n is the number of parameters in Ω . We can also add together (with suitable weights) several functions, in order to reduce the number of functions to be minimized (and the complexity of the corresponding multiobjective problem).

Next, to solve this multiobjective optimization problem, the algorithm called *Weighting Achievement Scalarizing Function Genetic Algorithm* (WASF-GA) is applied [16, 17, 52]. We choose this particular multiobjective algorithm as it has been applied with success in a previous work to identify some parameters of an epidemiological model simulating the evolution of Ebola Virus Disease outbreaks in Democratic Republic of the Congo [17].

From a general point of view, WASF-GA is an evolutionary algorithm that manages a population of possible solutions (called individuals), trying to improve them iteratively towards the optimal set (called Pareto optimal set). To do that, it uses two procedures at each iteration: reproduction and replacement. Firstly, the reproduction method applies some classical genetic operators (here: selection, mutation, and crossover) to the population of the previous iteration to obtain a new offspring population. Secondly, the replacement procedure decides which solutions from the previous iteration and the offspring populations should compose the population for the next iteration. The main feature of WASF-GA is that this decision is based on preferences, which are defined as reference values for the objectives. Indeed, it classifies the individuals using an achievement scalarizing function (see [52]) characterized by the L_∞ distance between the objective function values and the reference values with different weights. The considered sample of weight vectors is built following [52] to ensure they are well distributed.

As a result, WASF-GA returns a set of solutions covering a region of the Pareto optimal front.

In this work, we set the number of iterations to 15000, the population size to 50 individuals, the number of weight vectors to 50, and the reference values for the objectives to zero, corresponding to the ideal situation of a total matching between the model outputs and the reported data. The remaining input parameters of WASF-GA are those reported in [52]. Those values were set experimentally in order to obtain a good ratio between computational time and precision of the results.

Since WASF-GA is a stochastic algorithm, it may return different Pareto front approximations for different executions of the same instance. Then, to guarantee that results do not depend on the stochasticity of the algorithm, we perform 60 runs of each experiment. Among all the 60 sets of solutions, we extract the most balanced solution, meaning the one whose objective values are the closest to the reference values (in our case, zero) when considering the Euclidean distance.

In the 8th, 9th, 10th, 11th and 12th equations of system 1, which, as said previously, are not coupled with the other equations, there is a parameter, namely γ_Q , that only appears in those equations. Therefore, its estimation can be done once the multiobjective process is over. For instance, if the authorities report the time series corresponding to the number of people in quarantine at home (i.e. q_r), the parameter γ_Q can be estimated separately as a final step of the identification methodology. To do that, we assume that γ_Q can be parameterized by $n' \in \mathbb{N}$ parameters denoted by $\Phi = \{\Phi_1, \dots, \Phi_{n'}\} \in S' \subset \mathbb{R}^{n'}$ (see next section for a particular example). Then, we formulate the following mono-objective problem:

$$\begin{cases} \min f_q(\Phi) \\ \text{subject to } \Phi \in S', \end{cases} \quad (7)$$

where $f_q(\Phi)$ measures the error (as defined by (5)) between the reported data q_r and q_m^Φ , the number of people in quarantine returned by the model when considering γ_Q described by Φ .

Here, we solve this mono-objective problem by using a genetic algorithm configured with the same selection, mutation, and crossover operators as WASF-GA and the same input parameters (i.e. we set the number of iterations to 15000 and the population size to 50 individuals). Furthermore, we also perform 30 runs to guarantee that the results do not depend on the stochasticity of the algorithm.

4 Numerical simulations and calibration of the model for the case of Italy

We present here the results obtained with our model for the case of Italy, as an example of application. The situation in that country, like in other European countries such as Spain, the United Kingdom and France, has been especially complicated since it has been strongly affected by the disease. The first two known cases of Coronavirus in Italy, a couple of Chinese tourists, were confirmed on January 30. The first known case of secondary transmission occurred in Codogno, Municipality of Lombardy, on February 18, 2020. After that, the pandemic grew and extended to other countries, which needed the establishment of strict control measures (see [42]).

In Section 4.1, we introduce the data and some model parameters specific to the case of Italy. Then, in Section 4.2, we show and analyze some results returned after the numerical experiments performed to illustrate the efficiency of our approach.

4.1 Considered data and model parameters

The information regarding the first cases in Italy is available in [42]. More precisely, the first two known cases of Coronavirus in Italy, a couple of Chinese tourists, were confirmed on January 30. On February 26, they were declared recovered. The first known case of secondary transmission occurred in Codogno, Municipality of Lombardy, on February 18, 2020. Thus, in the numerical simulations, we assume that the index case started his incubation period in Italy on $t_0=19$ January 2020, as the earliest symptom onset of confirmed patients can be traced back to 30 January 2020 (reported on 31 January 2020) and we assume a delay of $d_I + d_E=10.5$ days (rounded to 11 days) between the beginning of the incubation period and the symptoms onset. Thus, we set, $E(t_0) = 1$, $S(t_0) = N - 1$ and all other compartments are set to 0.

The different control measures implemented in Italy are detailed in [42]. They started on $\lambda_1 = 23$ February 2020, where a 'red zone' was instituted on 11 municipalities with local transmission of SARS-CoV-2 infection, with lockdown to contain the emerging threat. Before this date we set $m_0 = m_1 = 1$ and after this date we consider $m_2 \in [0, 1]$ (to be estimated). On $\lambda_2 = 11$ March 2020, lockdown was declared on the entire country. Thus, we assume $m_3 = c_3 \times m_2$ with $c_3 \in [0, 1]$. On $\lambda_3 = 22$ March 2020 new additional measures regarding the containment and management of the epidemiological emergency from COVID-19, applicable on the whole national territory were adopted including the closure of non-essential or strategic production activities until May 3 (after several extensions). Consequently, we set $m_4 = 0$, corresponding to the most restrictive social distancing control measures. Starting on $\lambda_4 = 4$ May 2020, some production activities are authorized and we assume $m_5 = c_5 \times m_2$ with $c_5 \in [0, 1]$. The Decree of 16 May 2020 establishes the end of regional travel restrictions and the reopening of production activities from $\lambda_5 = 18$ May 2020. We set $m_6 = m_5 + c_6 \times (1 - m_5)$ with $c_6 \in [0, 1]$. Finally, this Decree also governs the end of total travel restrictions (traveling to and from abroad can only be limited with measures derived from the legal system of the European Union and international obligations) and the reopening of production, commercial and social activities from $\lambda_6 = 3$ June to 31 July ([43]). Hence, we consider $m_7 = m_6 + c_7 \times (1 - m_6)$ with $c_7 \in [0, 1]$.

For Italy we do not include the compartment I_{D_u} (i.e. people who die undetected pass from compartment I to compartment D_u). This simplification is done since, according to [31], almost all undetected deaths affect to vulnerable persons that die quickly after developing clinical sign (e.g. aged persons in nursing homes). All the other compartments were used in this case study.

Furthermore, according to [26], noticeable changes in the delay of detection of infected persons (i.e. d_I) have not been observed during the course of the epidemic in Italy. Thus, we set $d_g = 0$.

The information by regions concerning the number of cases (i.e. people who enter each day in compartments H_R or H_D) and COVID-19 deaths per day is daily reported in the website of the Italian Ministry of Health, (see [44]). Moreover, the Italian National Institute of Health supervises daily the data related to cases, deaths, tests carried out, etc. and the Italian civil protection publishes them completely here:

<https://github.com/pcm-dpc/COVID-19>

In our simulations, we take $c_r(t)$, $d_r(t)$, $\text{Hos}_r(t)$, $q_r(t)$ from the previous github repository, $h_r(t)$ from [28] and $\tau_{1,r}(t)$, $\tau_{2,r}(t)$ from [62].

We use the multiobjective technique described above to identify the parameters

$$\Omega = (\beta_{I,0}, \kappa_2, \kappa_3, \kappa_4, \kappa_5, \kappa_6, \kappa_7, c_2, c_3, c_5, c_6, c_7, p_0, C_u, \omega_{u,0}, 1/\gamma_{H_R}) \in \mathbb{R}^{16},$$

taking $\rho_1 = 10$ and $\rho_2 = 100$. We reduce the number of objective functions to two functions, $f_1(\Omega)$ and $f_2(\Omega)$ that combine the four time-series errors $f_c(\Omega)$, $f_d(\Omega)$, $f_{Hos}(\Omega)$ and $f_{d_u}(\Omega)$, using some weights as follows:

$$f_1(\Omega) = 0.5f_c(\Omega) + 0.5f_{Hos}(\Omega) \quad \text{and} \quad f_2(\Omega) = 0.9f_d(\Omega) + 0.1f_{d_u}(\Omega). \quad (8)$$

Thus, we solve the following two-objective problem:

$$\begin{cases} \min \{f_1(\Omega), f_2(\Omega)\} \\ \text{subject to } \Omega \in S \subset \mathbb{R}^{16}, \end{cases} \quad (9)$$

instead of solving a four objective one, that is more complex and time-consuming. We use $\rho_1 = 10$ (penalization to avoid the underestimation of the reported data in $[t_0, t_0 + T_{\max})$) and $\rho_2 = 100$ (penalization to avoid the underestimation of the reported data at $t_0 + T_{\max}$).

Regarding the data denoted by $q_r(t)$, reported by the Italian authorities, after several numerical experiments, we decided to define $\gamma_Q(t)$ as follows:

$$\gamma_Q(t) = \begin{cases} \gamma_{Q,1} & , \text{ if } t \leq t_{\gamma_Q}, \\ \gamma_{Q,2} & , \text{ if } t > t_{\gamma_Q}. \end{cases}$$

Next, we solve the mono-objective problem described by (7) to identify the parameters $\Phi = (\gamma_{Q,1}, \gamma_{Q,2}, t_{\gamma_Q}) \in [0, 1] \times [0, 1] \times [0, +\infty)$.

The effective reproduction number given by (2) has been computed using the Composite Simpson's rule.

4.2 Results

First, we use the methodology and techniques detailed in Sections 2 and 3 to identify suitable model parameters for the case of Italy, validating the model proposed here. Then, we show how the model may be used to estimate the spread of the disease for different futures scenarios.

4.2.1 Validation of the model.

There is much uncertainty about θ but, as explained in the Section 2.1, we can first estimate $\frac{\omega(t)}{\theta(t)}$ and then $\theta(t)$ for different values of ω (which is the instantaneous infection detected fatality ratio, iIdFR). That shows that, actually, the uncertainty of θ is related to the uncertainty of ω . Hence, we have considered different scenarios for the value of ω , using several estimations found in the literature for the Infection Fatality Ratio (IFR). We remind that the IFR is the expected value of the probability of dying for a person who is infected, i.e. the number of deaths due to the virus (detected or not) over the number of infections (detected or not). The IFR is one of the most critical features of the coronavirus disease, being quite difficult to estimate. More precisely:

- According to [55], published on 4 May 2020, IFR=1.2%. Then, since on 5 May 2020 there were 29315 detected deaths (see [13]) and 9154 estimated undetected deaths (see [31, 32]), an associated estimation of ω would be $\omega = 1.2\% \frac{29315}{29315+9154} = 0.9145\%$.
- According to [21], published on 18 May 2020, IFR=1.91%. Then, an associated estimation of ω would be $\omega = 1.91\% \frac{29315}{29315+9154} = 1.4555\%$.
- According to [11], published on 21 May 2020 but received on 4 April 2020, IFR=5.7%. Then, since on 4 April 2020 there were 15362 detected deaths and on 14 April 2020 there were 6773 estimated undetected deaths (see [30, 32]), an associated estimation of ω would be $\omega = 5.7\% \frac{15362}{15362+6773} = 3.9559\%$.

We have used these three values of ω up to 21 July 2020.

The whole set of model parameters obtained when considering $\omega = 1.4555\%$ is reported in Tables 1 and 2. For the sake of simplicity, we only present here the parameters obtained with this particular value. Results for other values ω can be found at <https://github.com/momat-ucm/T-SIR-T>.

In Figures 2 and 3, we show the comparison of some of the outputs of the simulation run with the chosen parameters and the official real data, when $\omega = 1.4555\%$. Similar figures have been obtained when considering $\omega = 0.9145\%$ and 3.9559% (obviously, with different parameters) and, thus, are not reported here.

More precisely, in Figure 2, we show a very good fitting of the model outcomes with respect to four important time series reported by the Italian authorities regarding the pandemic: the cumulative number of cases and deaths, people in hospital, people in quarantine and recovered people. Furthermore, the model estimates that on 14 April 2020 and 5 May 2020 there were around 7509 and 9192 undetected deaths (persons in D_u), respectively. According to the Italian authorities (see [32, 31]), on 14 April 2020 and 5 May 2020 (the only two reported values) this number was estimated to be 6773 and 9154, respectively. This shows that our model is able to estimate well the biological behavior of the pandemic.

Next, in Figure 3 (Top-Left) and (Top-Right), we show the evolution of the new detected cases and deaths, per day. When available, we compare those outputs with the real data. Again we observe a good fitting between simulated results and observed data. This feature of the model is interesting as it allows to estimate the date and magnitude of the peaks of all those outcomes (also reported in this figure), which may help to plan some health system measures (such as increasing the number of beds in hospital).

In Figure 3 (Bottom-Left) and (Bottom-Right), we show the evolution of the functions corresponding to the social distancing measures and the effective reproduction number, when $\omega = 1.4555\%$. We observe that the effective reproduction number went below 1 after 14 March 2020, which should indicate the start of a decrease of the spread in Italy. This value is reached soon after the implementation of some intensive control measures (lockdown in the entire country) in Italy on 11 March 2020, showing that such restrictive control measures were necessary to control the pandemic (before this date the model estimates that R_e remains above 1). We also note that the initial value of the effective reproduction number, which is the basic reproduction number R_0 , is equal to 8.0, which is a value higher than those generally reported in the literature (between 1.4 and 3.9, see e.g. [8, 39]). However, we also show in this figure the contribution of each infectious state (i.e. E , I , I_u , H_R and H_D) to the value of R_e (see Section 2.2 for more details). In particular, we observe that the compartment I_u of undetected infectious people contributes to R_0 with a value of 3.9, which means that detected cases contribute to R_0 with a value of less than 4.1, which is consistent with the values usually found in the literature.

Next, in Figure 4, we show the evolution of some of the outcomes of the model for different values of ω . First, in this figure (Top-Left) and (Top-Right), we present the evolution of R_e and the effect of the social distancing measures when considering $\omega = 0.9145\%$, $\omega = 1.4555\%$ and 3.9559% . We observe that, before 11 March 2020, the magnitude and shape of those functions are affected by the value of ω . After this date, they are quite similar for the different values of ω . Focusing on the total number of cases (including undetected), presented in Figure 4 (Bottom-Left), the model estimates that, taking into account the different values of ω considered previously, between 1 and 4 million people could have been infected by COVID-19 in Italy. Regarding the value of θ , depicted in Figure 4 (Bottom-Right), we see that for $\omega = 3.9559\%$ the proportion of detection of new cases oscillates between 15% and 67%. For $\omega = 0.9145\%$ or $\omega = 1.4555\%$, it is between 5.7% and 25%. From those results we see that, in all cases, the model estimates that a large amount of new cases have not been detected during this pandemic. However, the proportion of detection is highly dependent on the value of ω and large variations of the results are observed when this value changes. The results presented in Figure 4, indicate the necessity to approximate well this critical parameter.

4.2.2 Estimation of possible future scenarios

We considered just a few cases, to illustrate the possibilities that can be explored with the model to estimate the possible spread of the disease. We simulate the future scenarios described below until 1 September 2020. We recall that the meaning of parameters λ_i , κ_i and m_i that appear below may be seen in Section 2.1. In all cases, we consider $\lambda_7 = 22$ July 2020 (date of a potential change of control measures; taking into account that the date of our last data is 21 July 2020), $\kappa_8 = \kappa_2 = 62.0015$. Since we cannot estimate θ for the future (it depends on the authority actions)

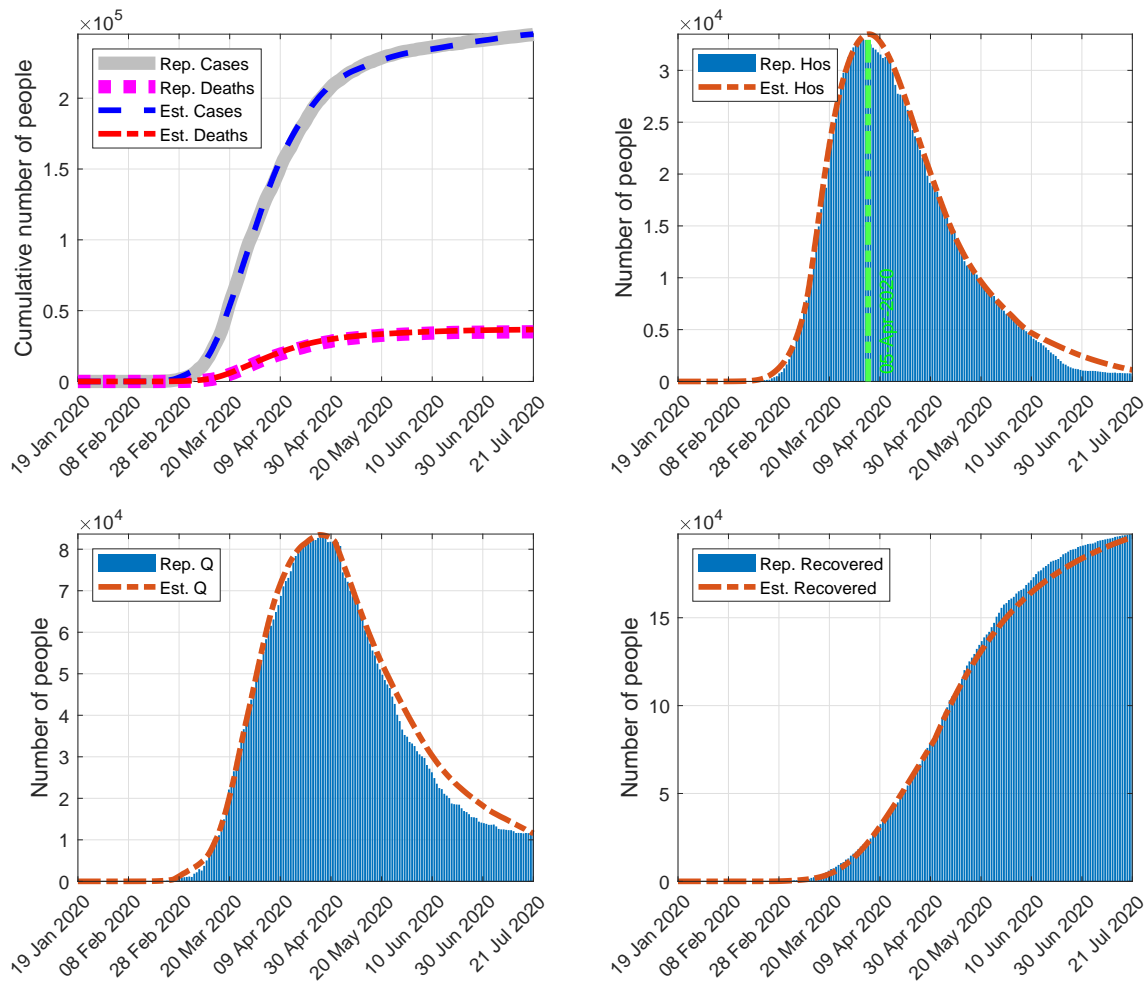


Figure 2: Comparison of some of the outputs of the simulation run with $\omega = 1.4555\%$ (Est.) and the Italian official reported data (Rep.). Top-left: Cumulative number of cases and deaths. Top-right: People in hospital. Bottom-left: People in quarantine. Bottom-right: Cumulative number of recovered people.

Table 1: Summary of all parameters used for COVID-19 in Italy when considering $\omega = 1.4555\%$. A brief description (**Description**) of each parameter is given. Furthermore, if available, the range of the considered values (**Value**) and the reference in the literature (**Ref.**) are reported. The notation * means that this value is identified during the multiobjective or monoobjective optimization processes. The notation – means this parameter is omitted during this work. The notation & means that this parameter (depending on time) is estimated directly from the reported data. The notation \diamond means that the value of the parameter is set by the authors of this article, after some numerical experiments.

Notation	Value	Description	Ref.
N	60317000	Number of people in the country before starting the pandemic	[59]
γ_E	$\frac{1}{5.5}$	Transition rate of a person in compartment E (day^{-1})	[64]
$\gamma_I(t)$	$\frac{1}{5}$	Transition rate of a person in compartment I (day^{-1}) at time t	[25]
$\gamma_{I_u}(t)$	$\frac{1}{9}$	Transition rate of a person in compartment I_u to compartment R_u (day^{-1}) at time t	[25]
$\gamma_{H_R}(t)$	$\frac{1}{14.2729}$	Transition rate of a person in compartment H_R to compartment Q (day^{-1}) at time t	*
$\gamma_{H_D}(t)$	$\frac{1}{5}$	Transition rate of a person in compartment H_D to compartment D (day^{-1}) at time t	[25]
$\gamma_{I_{D_u}}(t)$	∞	Transition rate of a person in compartment I_{D_u} to compartment D_u (day^{-1}) at time t	[31]
$\gamma_{Q,1}$	$\frac{1}{36.0450}$	Transition rate of a person in compartment Q to compartment R_d (day^{-1}) at time $t \leq t_{\gamma_Q}$	*
$\gamma_{Q,2}$	$\frac{1}{24.8646}$	Transition rate of a person in compartment Q to compartment R_d (day^{-1}) at time $t > t_{\gamma_Q}$	*
t_{γ_Q}	2 May 2020	Date when the duration of a person in quarantine changes	*
d_g	0	Maximum number of days that d_I can be decreased	[26]
C_E	0.3806	Ratio between the disease contact rates β_E and β_I	[33]
C_u	0.3293	Ratio between $\underline{\beta}_I$ and the disease contact rate β_I	*
$C_{I_{D_u}}$	–	Ratio between the disease contact rates $\beta_{I_{D_u}}$ and β_I	–
$\beta_{I,0}$	0.4992	Disease contact rate of a person in compartment I without control measures (day^{-1})	*
$\beta_{I_{D_u},0}$	–	Disease contact rate of a person in compartment I_{D_u} without control measures (day^{-1})	–
β_{H_R}	&	Disease contact rate of a person in compartment H_R	&
β_{H_D}	&	Disease contact rate of a person in compartment H_D	&
t_0	19 Jan 2020	Initial time	[42]
$\underline{\omega}$	1.4555	iIdFR when the implemented control measures are fully applied (in %)	[21]
δ_ω	0	to compute iIdFR when no control measures are applied $\bar{\omega} = \underline{\omega} + \delta_\omega$	\diamond
$\omega_{u,0}$	0.42	iIuFR when $t = t_0$ (in %)	*
$\omega_{u,1}$	0.42	iIuFR when $t = t_{u,1}$ (in %)	\diamond
$\omega_{u,2}$	0	iIuFR when $t = t_{u,2}$ (in %)	\diamond
$t_{u,1}$	3 May 2020	One day before the last report of $d_{r,u}$	[44]
$t_{u,2}$	4 May 2020	Date of the last report of $d_{r,u}$	[44]
$\theta(t)$	&	Proportion of infected people that are detected by the authorities	&
p_0	0.7382	Ratio of the number of new detected infected people that will survive the disease and are hospitalized at time t_{θ_0}	[44]

Table 2: Summary of all parameters related to control measures used for COVID-19 in Italy when considering $\omega = 1.4555\%$. A brief description (**Description**) of each parameter is given. Furthermore, if available, the range of the considered values (**Value**) and the reference in the literature (**Ref.**) are reported. The notation * means that this value is optimized during the multiobjective optimization process. The notation – means this parameter is omitted during this work.

Notation	Value	Description	Ref.
λ_0	19 Jan 2020	First day of application of the control strategy that was being used before t_0 , if any	[42]
λ_1	23 Feb 2020	First day of application of the 1st control strategy (lockdown on 11 municipalities)	[42]
λ_2	11 Mar 2020	First day of application of the 2nd social distancing strategy (lockdown on the entire country)	[42]
λ_3	22 Mar 2020	First day of application of the 3rd social distancing strategy (closure of non-essential or strategic production activities)	[42]
λ_4	4 May 2020	First day of application of the 4th social distancing strategy (end of closure of non-essential or strategic production activities)	[42]
λ_5	18 May 2020	First day of application of the 5th social distancing strategy (end of regional travel restrictions and reopening of production activities)	[42]
λ_6	3 Jun 2020	First day of application of the 6th social distancing strategy (progressive end of total travel restrictions, reopening of production, commercial and social activities)	[43]
m_0	1	Intensity of the social distancing strategy used before t_0	[42]
m_1	1	Intensity of the control strategy that was being used in $(t_0, \lambda_1]$	[42]
m_2	0.5332	Intensity of the social distancing strategy used in $(\lambda_1, \lambda_2]$	*
m_3	0.1369	Intensity of the social distancing strategy used in $(\lambda_2, \lambda_3]$	*
m_4	0	Intensity of the social distancing strategy used in $(\lambda_3, \lambda_4]$	[33]
m_5	0.0549	Intensity of the social distancing strategy used in $(\lambda_4, \lambda_5]$	*
m_6	0.0577	Intensity of the social distancing strategy used in $(\lambda_5, \lambda_6]$	*
m_7	0.0578	Intensity of the social distancing strategy used in $(\lambda_6, \lambda_7]$	*
$m^{(t)}$	–	Intensity of the contact tracing control strategy	–
$m^{(s)}$	–	Intensity of the health system control strategy	–
κ_1	–	Efficiency of the social distancing strategy used in $(\lambda_0, \lambda_1]$	–
κ_2	62.0015	Efficiency of the social distancing strategy used in $(\lambda_1, \lambda_2]$	*
κ_3	0.5139	Efficiency of the social distancing strategy used in $(\lambda_2, \lambda_3]$	*
κ_4	0.0362	Efficiency of the social distancing strategy used in $(\lambda_3, \lambda_4]$	*
κ_5	90.6408	Efficiency of the social distancing strategy used in $(\lambda_4, \lambda_5]$	*
κ_6	14.6514	Efficiency of the social distancing strategy used in $(\lambda_5, \lambda_6]$	*
κ_7	52.0747	Efficiency of the social distancing strategy used in $(\lambda_6, \lambda_7]$	*
$\kappa^{(t)}$	–	Efficiency of the contact tracing control strategy	–
$\kappa^{(s)}$	–	Efficiency of the health system control strategy	–

by using the methodology described in Section 3.2, from 22 July 2020 on, we assume that θ remains equal to its value on 21 July 2020.

First, for $\omega = 0.9145\%$, $\omega = 1.4555\%$ and $\omega = 3.9559\%$, we consider $m_8 = 0.25$. Some results of this simulation are reported in Figure 5 (Top-Left). In the three cases we observe a slight increase in the number of cases.

Next, for $\omega = 1.4555\%$, we simulate four future scenarios with: (i) $m_8 = m_7 = 0.0578$; (ii) $m_8 = 0.15$; (iii) $m_8 = 0.25$, and (iv) $m_8 = 0.4$. In Figure 5 (Top-Right), (Bottom-Left) and (Bottom-Right), we show some outputs

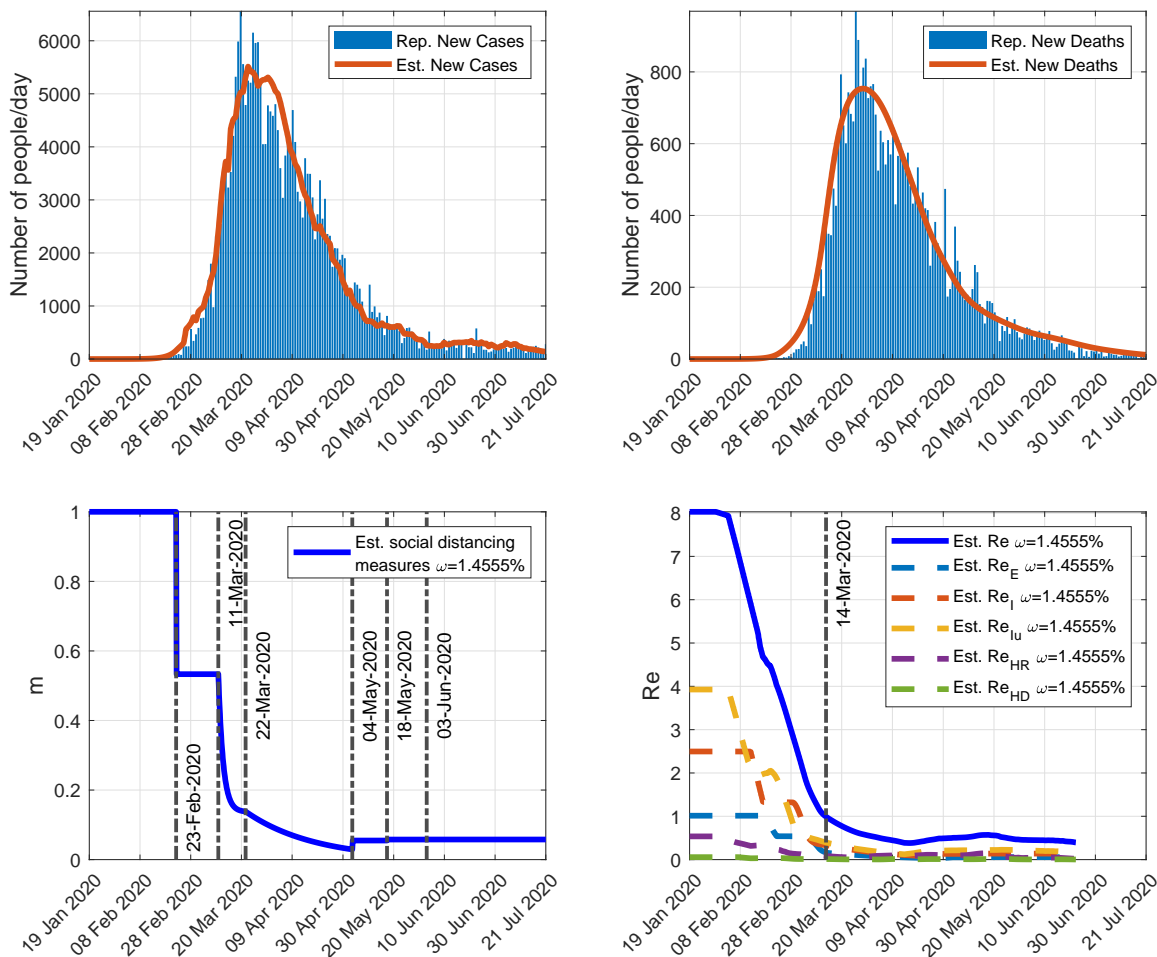


Figure 3: Comparison of some of the outputs of the simulation run with $\omega = 1.4555\%$ (Est.) and, when available, the Italian official reported data (Rep.). Top-left: New detected cases per day. Top-right: New detected deaths per day. Bottom-Left: Function modeling the social distancing measures. Bottom-Right: Effective reproduction number and the contribution of each infectious state (i.e. E , I , I_u , H_R and H_D) to this number. The vertical line corresponds to the first date when $R_e < 1$.

of the model for these four scenarios. As we can see, there is a remarkable possibility of new outbreaks in Italy if the control measures are relaxed too much. For instance, focusing on the effective reproduction number, it goes back over 1 when $m_8 = 0.25$ and $m_8 = 0.4$, which would mean a second wave of uncontrolled outbreaks. In particular, if control measures are relaxed too much (e.g. $m_8 = 0.4$) the new cases and deaths may increase dramatically. This clearly indicates that, the pandemic is not over yet and efforts must continue, keeping enough control measures.

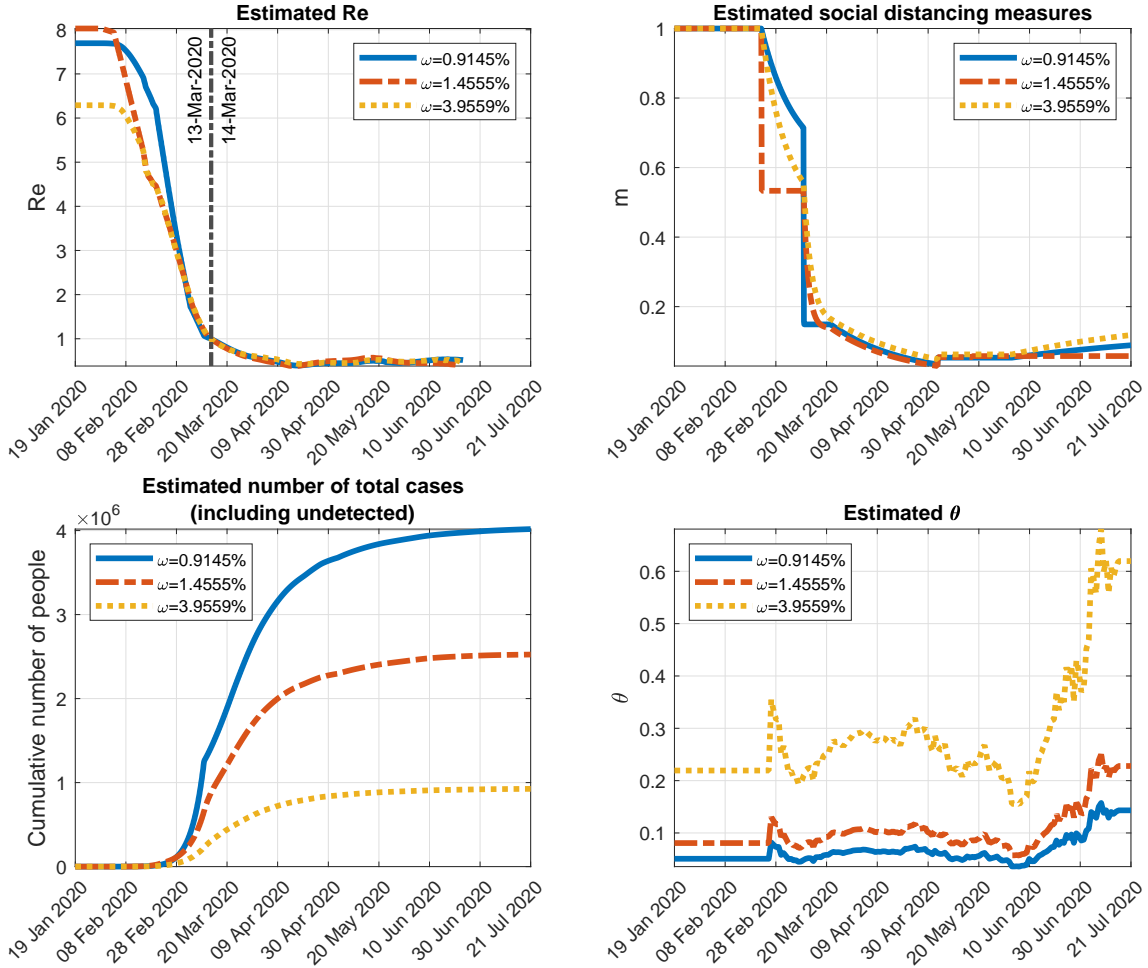


Figure 4: Some of the outputs of the simulation run with $\omega = 1.4555\%$, $\omega = 0.9145\%$ and 3.9559% for the Italian case. Top-left: Effective reproduction number. Top-right: Function modeling the social distancing measures. Bottom-left: Cumulative number of total cases (including undetected cases). Bottom-right: $\theta(t)$.

5 Conclusions and Discussion

We have developed an epidemiological model for simulating the dynamics of COVID-19 in a considered territory. This model may be used by epidemiologists and policy makers for understanding the spread of the pandemic, study the efficiency of control measures and propose suitable ways to tackle the problem. It is intended to be as simple as possible, but with sufficient complexity to be able to simulate the biological and sociological mechanisms which influence the disease spread (including the three main types of control measures: social distancing, contact tracing and health system measures), to fit real data and to simulate possible future scenarios. Of course, for particular cases the model may be simplified (for instance removing the I_{D_u} compartment, as we showed for the case of Italy, or the Q compartment, if the territory does not keep asymptomatic people in quarantine at home).

The model can serve as a basis for studies including higher complexity. Indeed, it can be modified to introduce

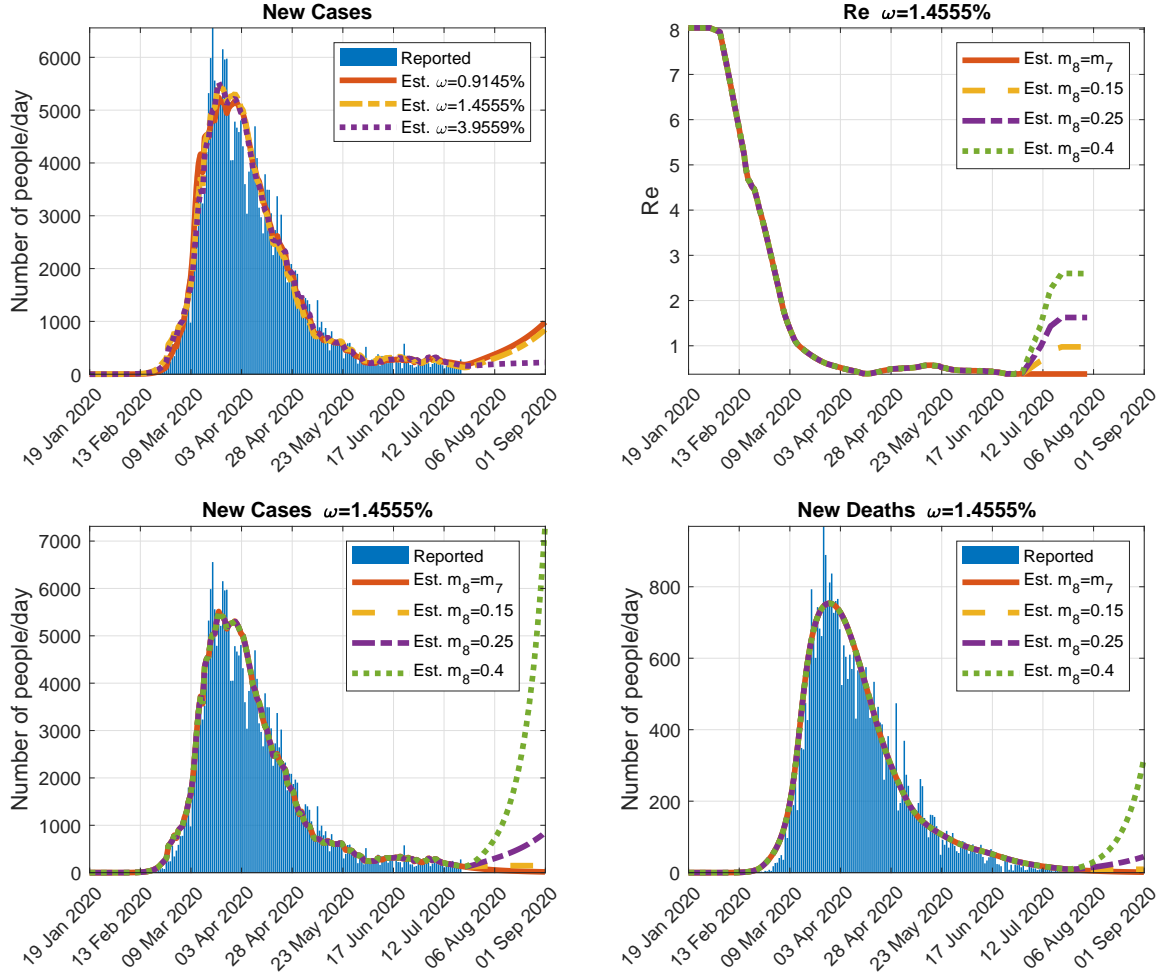


Figure 5: Outputs obtained when simulating possible future scenarios (using $\kappa_8 = \kappa_2 = 62.0015$). Top-left: New detected people when $m_8 = 0.25$ and $\omega = 0.9145\%$, 1.4555% or 3.9559% , Top-right: Effective reproduction number when $\omega = 1.4555\%$ and $m_8 = 0.15, 0.25$ or 0.4 , Bottom-left: New detected cases per day, when $\omega = 1.4555\%$ and $m_8 = 0.15, 0.25$ or 0.4 , Bottom-right: New detected deaths per day, when $\omega = 1.4555\%$ and $m_8 = 0.15, 0.25$ or 0.4 .

structures of age, sex, etc., stochasticity, levels of uncertainty, mobility between territories, spatial structure, etc.

In the model that we propose, there are several important points to highlight that, to the best of our knowledge, have not been previously considered:

1. We have included a compartment for undetected deaths mainly happening in nursing homes or residences for disabled people. The number of these deaths and the associated infections does not seem to be negligible, which makes necessary to take them into account in the model.
2. Some dynamic coefficients are based on reported time series, which allows the model to reproduce faithfully the spread of the disease, fitting very well the reported data.
3. The model includes the effects of the different control measures described above.

4. The model takes into account quarantine and hospitalized people in order to estimate the number of hospital beds that are needed.

We point out that the model, in this current simple form, has the following limitations:

1. Epidemics/pandemics are phenomena that respond to complex processes and all epidemiological models are simplifications of this complex reality. Furthermore, the model is based on the interactions of human beings, which do not behave in reasonably predictable ways like molecules, cells or particles (see [22]).
2. The prediction of its evolution is very complicated because of the characteristics of the system of differential equations governing the model. In particular, in the exponential growth phases, small variations in some parameters can lead to large variations in model outcomes (see, [33]). Moreover, there are data quality problems: a) uncertainty about the characteristics of SARS-CoV-2 (new virus); b) changes in criteria to count cases; c) uncertainty about undetected cases (and, therefore, the mortality rate), etc.
3. Since there is no clear scientific evidence of the effect of humidity and temperature (see, e.g. [40, 56]), we have not included these two factors in our model (if such evidence is demonstrated, these variables should be included in the model).
4. It is assumed that the spatial distribution of the population is homogeneous in the territory under study (if this is not a suitable assumption, the territory can be divided into smaller ones).
5. In this paper we have only considered the spread of disease within a territory, where only local transmission and known imported / exported cases are modeled. The between-country spread has not been modeled in this work, but can be considered in the future employing ideas from the BeCoDiS model (see [35]).

Acknowledgments

This research was partially supported by the Spanish Government under projects PID2019-106337GB-I00, MTM2015-64865-P and MTM2017-89423-P; the research group MOMAT (Ref. 910480) supported by the Complutense University of Madrid; and the Regional Government of Andalusia under project P12-TIC301, partially financed by the European Regional Development Fund (ERDF). Also thanks to M. Ramos-Rubio for his help with the manuscript.

Annex

In the version of this article initially submitted to *Physica D: Nonlinear Phenomena*, the case of Italy studied in Section 4 was written using official data up to 21 July 2020 to validate the model and analyze several scenarios up to 1 September 2020. In the current version, written later as a revised version of the submitted paper, we have kept everything that was computed using those data (including the values of the parameters). Furthermore, following the reviewers' suggestions, in this Annex we show an update of the situation in Italy, using official data up 21 December 2020, to compare with our simulations. It is important to remark that in all the simulations shown in this Annex, we keep the same parameter values as those given in the version of the paper initially submitted to the journal (with data up to 21 July 2020, as mentioned above).

In all cases, we consider $\omega = 1.4555\%$ and we estimate θ by using the methodology described in Section 3.2 until 21 December 2020 assuming that, since then, θ remains equal to its value on this day. Between 14 and 15 August 2020, the cumulative number of deaths reported by the Italian authorities shows a sudden increase, because they updated unreported deaths that occurred in previous weeks. According to their weekly reports (see [27]), from 10 to 23 August there were 14 new deaths. Thus, 218 deaths must be smoothly distributed before the 10 August as done in [33] for the case of China. First, we distribute the 14 deaths from 10 to 23 August and, then, we assign the 218 deaths between 22 July and 23 August, avoiding the modification of the data that we had already used in the paper until 21 July.

Besides the values of the model parameters shown in Section 4, in Table 3 we show the values of the parameters that we have used to model the control measures (as implemented for the Italian Government) corresponding to the days after the 21 July 2020.

Table 3: Summary of the parameters used to model the control measures corresponding to the days after the 21 July 2020. Time intervals with different measures are indexed with $i \in \{8, 9, 10, 11, 12, 13\}$.

Date of application (λ_{i-1})	Efficiency (κ_i)	Intensity (m_i)	Type	Detail
1 Aug 2020	62.0015	0.3	Relaxation	“New normality”
20 Sep 2020	62.0015	0.53	Relaxation	Back from vacation, school, work. Regional elections
3 Nov 2020	62.0015	0.3	Restriction	Curfew
13 Nov 2020	62.0015	0.22	Restriction	More regions with restrictive measures
23 Dec 2020	62.0015	Different scenarios	Different scenarios	Christmas
7 Jan 2020	62.0015	Different scenarios	Different scenarios	End of Christmas

In Figure 6 we show some results corresponding to five different scenarios up to 31 January 2021:

- Red scenario: More restrictive control measures are implemented from 23 December to 31 January.
- Yellow scenario: More restrictive control measures are implemented from 23 December to 6 January. From 7 January, the measures come back to the ones before 23 December.
- Purple scenario: The control measures are maintained without changes.
- Green scenario: The control measures are relaxed for Christmas from 23 December to 6 January. From 7 January, more restrictive control measures are implemented.
- Cyan scenario: The control measures are relaxed for Christmas from 23 December. These measures continue during January, i.e., no restrictions are implemented after Christmas.

As we can see, Italy could be on 22 December at a crucial moment where any changes in control measures could have notable consequences. The model estimates that on 22 December the effective reproduction number R_e could be slightly higher than 1 and the intensity of the control measures could be of 0.22 (considering 1 the absence of measures before the pandemic). If they tighten the measures from 23 December, as the Italian government has announced, the effective reproduction number could drop below 1 and the pandemic would be controlled (see the red scenario in Figure 6). However, if the measures are relaxed too soon in January, the cases and deaths would increase again before the end of January (see the yellow scenario in Figure 6). As we can see in the purple scenario, if no changes are implemented in the measures from 22 December to 31 January, the new cases and deaths may increase. Finally, if the measures are relaxed on 23 December, there would be a third wave of uncontrolled outbreaks whose magnitude depends on the restrictive measures they apply after Christmas (see the green and cyan scenarios in Figure 6).

References

- [1] D. Adam. A guide to R — the pandemic’s misunderstood metric. *Nature, News Feature*, 3 July 2020. doi: <https://doi.org/10.1038/d41586-020-02009-w>.
- [2] M. M. Arons, K. M. Hatfield, S. C. Reddy, A. Kimball et al. Presymptomatic SARS-CoV-2 infections and transmission in a skilled nursing facility. *New England Journal of Medicine*, 382(22):2081–2090, 2020. doi: 10.1056/NEJMoa2008457.

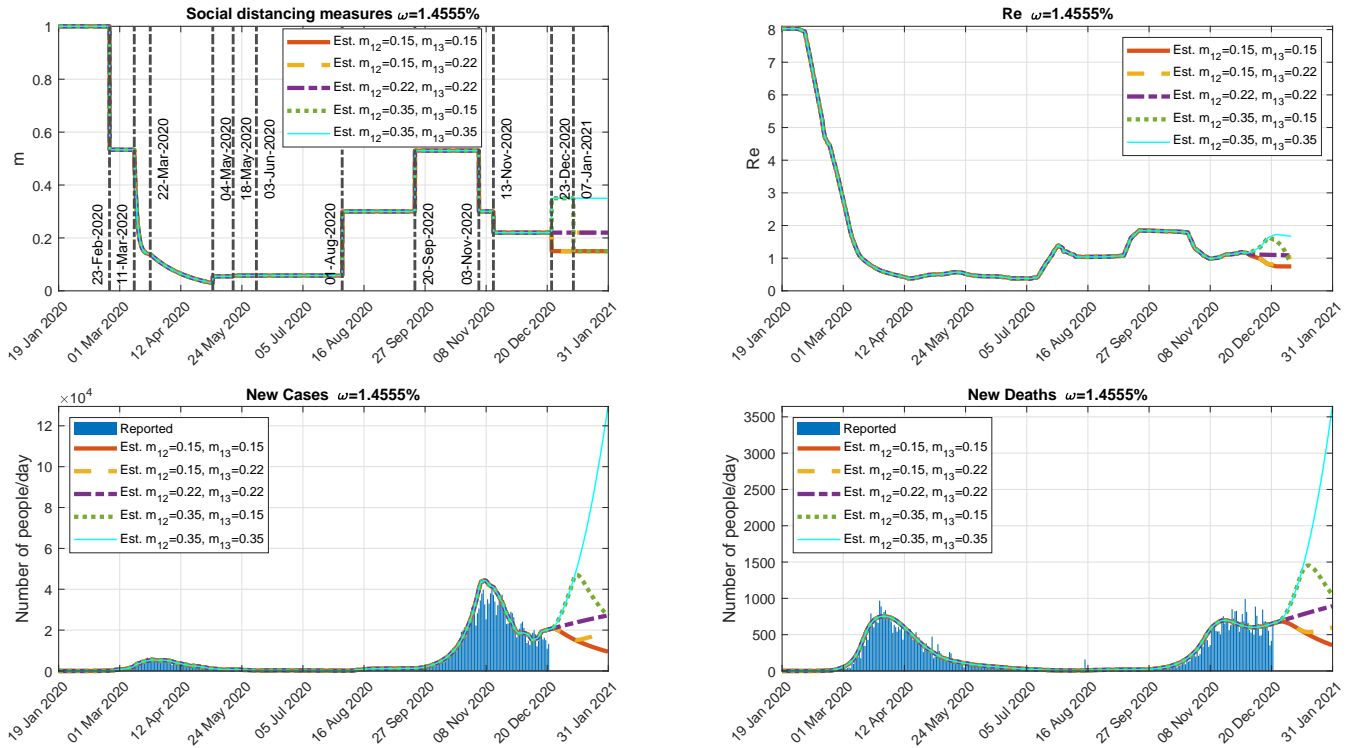


Figure 6: Outputs obtained when simulating possible future scenarios when $\omega = 1.4555\%$. Top-left: Function modeling the social distancing measures. Top-right: Effective reproduction number. Bottom-left: New detected cases per day. Bottom-right: New detected deaths per day.

[3] A. Basu. Estimating the infection fatality rate among symptomatic COVID-19 cases in the United States. *Health Affairs*, 39(7):1229–1236, 2020. <https://doi.org/10.1377/hlthaff.2020.00455>.

[4] E. Bertuzzo, L. Mari, D. Pasetto, S. Miccoli, R. Casagrandi, M. Gatto and A. Rinaldo. The geography of COVID-19 spread in Italy and implications for the relaxation of confinement measures. *Nature Communications*, 11:4264, 2020. doi:<https://doi.org/10.1038/s41467-020-18050-2>.

[5] F. Brauer and C. Castillo-Chávez. *Mathematical Models in Population Biology and Epidemiology*. Texts in Applied Mathematics, Springer, 2001. URL: <https://www.springer.com/gp/book/9781441931825>.

[6] S. Chen, J. Yang, W. Yang, and T. Barnighausen. COVID-19 control in China during mass population movements at New Year. *The Lancet*, 395(10226):764–766, 2020. doi:[https://doi.org/10.1016/S0140-6736\(20\)30421-9](https://doi.org/10.1016/S0140-6736(20)30421-9).

[7] H. Chu, J. Li, D. Blakeley, and R.J. Smith. The failure of R_0 . *Computational and Mathematical Methods in Medicine*, 527610, 2011. doi:10.1155/2011/527610.

[8] M. D’Arienzo and A. Coniglio. Assessment of the sars-cov-2 basic reproduction number, r_0 , based on the early phase of covid-19 outbreak in italy. *Biosafety and Health*, 2(2):57 – 59, 2020. doi:<https://doi.org/10.1016/j.bsheal.2020.03.004>.

- [9] Centers for Disease Control and Prevention. Transmission of SARS-COV-2 Infections in Households — Tennessee and Wisconsin. *Morbidity and Mortality Weekly Report*, 69(44):1631-1634, 2020. <https://www.cdc.gov/mmwr/volumes/69/wr/pdfs/mm6944e1-H.pdf>.
- [10] F. Della Rossa, D. Salzano, A. Di Meglio *et al.* A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic. *Nature Communications*, 11:5106, 2020. doi:<https://doi.org/10.1038/s41467-020-18827-5>.
- [11] G. De Natale, V. Ricciardi, G. De Luca, D. De Natale *et al.* The COVID-19 infection in Italy: A statistical study of an abnormally severe disease. *J. Clin. Med.*, 9(5):1564, 2020. doi:<https://doi.org/10.3390/jcm9051564>.
- [12] O. Diekmann, H. Heesterbeek, and T. Britton. *Understanding Infectious Disease Dynamics*. Princeton Series in Theoretical and Computational Biology, Princeton University Press, 2013. URL: <http://www.jstor.org/stable/j.cttq9530>.
- [13] Dipartimento della Protezione Civile. Aggiornamento casi COVID-19. URL: <http://opendatadpc.maps.arcgis.com/apps/opsdashboard/index.html#/b0c68bce2cce478eaac82fe38d4138b1>, 2020.
- [14] European Centre for Disease Prevention and Control. Surveillance of COVID-19 at long-term care facilities in the EU/EEA. URL: <https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-long-term-care-facilities-surveillance-guidance.pdf>, May 2020.
- [15] M. R. Ferrández, B. Ivorra, J. L. Redondo, A. M. Ramos, and P. M. Ortigosa. A multi-objective approach to estimate parameters of compartmental epidemiological models. Application to Ebola Virus Disease epidemics. *Preprint in ResearchGate*, 2020. doi:<https://doi.org/10.13140/RG.2.2.25778.56006>.
- [16] M. R. Ferrández, S. Puertas-Martín, J. L. Redondo, B. Ivorra, A. M. Ramos, and P. M. Ortigosa. High-performance computing for the optimization of high-pressure thermal treatments in food industry. *The Journal of Supercomputing*, 75:1187–1202, 2018. doi:10.1007/s11227-018-2351-4.
- [17] M. R. Ferrández, J. L. Redondo, B. Ivorra, A. M. Ramos, and P. M. Ortigosa. Preference-based multi-objectivization applied to decision support for High-Pressure Thermal processes in food treatment. *Applied Soft Computing*, 79:326–340, 2019. doi:<https://doi.org/10.1016/j.asoc.2019.03.050>.
- [18] M. Gatto, E. Bertuzzo, L. Mari, S. Miccoli, L. Carraro, R. Casagrandi and A. Rinaldo. Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *PNAS*, 117(19):10484-10491, 2020. doi:<https://doi.org/10.1073/pnas.2004978117>.
- [19] G. Giordano, F. Blanchini, and R. Bruno *et al.* Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*, 2020. doi:<https://doi.org/10.1038/s41591-020-0883-7>.
- [20] D.H. Glass. European and US lockdowns and second waves during the COVID-19 pandemic. *Mathematical Biosciences*, 330:108472, 2020. doi:<https://doi.org/10.1016/j.mbs.2020.108472>.
- [21] R. Grewelle and G. A. De Leo. Estimating the global infection fatality rate of COVID-19. *medRxiv*, 2020. doi:<https://doi.org/10.1101/2020.05.11.20098780>.
- [22] H.W. Hethcote. *The basic epidemiology models: models, expressions for R_0 , parameter estimation, and applications*. In *Mathematical Understanding of Infectious Disease Dynamics*, Chapter 1, pages 1–61. (eds.) S. Ma and Y. Xia. World Scientific, 2009. doi:10.1142/9789812834836_0001.
- [23] G. Huang, Y. Takeuchi, W. Ma and D. Wei. Global Stability for Delay SIR and SEIR Epidemic Models with Nonlinear Incidence Rate. *Bulletin of Mathematical Biology*, Volume 72, 1192–1207, 2010. doi:<https://doi.org/10.1007/s11538-009-9487-6>.

- [24] Instituto de Salud Carlos III. Vigilancia de los excesos de mortalidad por todas las causas. MoMo. URL: https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/MoMo/Documents/informesMoMo2020/MoMo_Situacion%20a%2021%20de%20diciembre_CNE.pdf, December 2020.
- [25] Istituto Superiore di Sanità. Characteristics of SARS-CoV-2 patients dying in Italy Report based on available data on May 21st, 2020. URL: https://www.epicentro.iss.it/en/coronavirus/bollettino/Report-COVID-2019_21_may_2020.pdf, May 2020.
- [26] Istituto Superiore di Sanità. Epidemia COVID-19, aggiornamento nazionale. URL: https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19_23-giugno-2020.pdf, June 2020.
- [27] Istituto Superiore di Sanità. Epidemia COVID-19, aggiornamento nazionale. URL: https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19_25-agosto-2020.pdf, August 2020.
- [28] Istituto Superiore di Sanità. Infografica. URL: <https://www.epicentro.iss.it>, 2020.
- [29] Istituto Superiore di Sanità. Recommendation for people in self-isolation and for the family members assisting them. 2020. <https://www.iss.it/documents/20126/0/Poster+Raccomandazioni+isolamento+rev+12+marzo+%281%29.pdf/d409de34-8bba-a220-10e8-e7e8a33963c3>.
- [30] Istituto Superiore di Sanità. Survey nazionale sul contagio COVID-19 nelle strutture residenziali e sociosanitarie. URL: <https://www.epicentro.iss.it/coronavirus/pdf/sars-cov-2-survey-rsa-rapporto-3.pdf>, April 2020.
- [31] Istituto Superiore di Sanità. Survey nazionale sul contagio COVID-19 nelle strutture residenziali e sociosanitarie. URL: <https://www.epicentro.iss.it/coronavirus/pdf/sars-cov-2-survey-rsa-rapporto-finale.pdf>, May 2020.
- [32] Istituto Nazionale di Statistica ed Istituto Superiore di Sanità. Impatto dell'epidemia COVID-19 sulla mortalità totale della popolazione residente primo quadrimestre 2020. URL: https://www.epicentro.iss.it/coronavirus/pdf/Rapp_Istat_Iss_3Giugno.pdf, June 2020.
- [33] B. Ivorra, M. R. Ferrández, M. Vela-Pérez, and A.M. Ramos. Mathematical modeling of the spread of the coronavirus disease 2019 (COVID-19) taking into account the undetected infections. the case of China. *Communications in Nonlinear Science and Numerical Simulation*, 88:105303, 2020. doi:<https://doi.org/10.1016/j.cnsns.2020.105303>.
- [34] B. Ivorra, D. Ngom, and A.M. Ramos. Stability and sensitivity analysis of Be-CoDiS, an epidemiological model to predict the spread of human diseases between countries. Validation with data from the 2014-16 West African Ebola Virus Disease epidemic *Electronic Journal of Differential Equations*, 2020, No. 62, pp. 1-29. <https://ejde.math.txstate.edu/Volumes/2020/62/ivorra.pdf>.
- [35] B. Ivorra, A. M. Ramos, and D. Ngom. Be-CoDiS: A mathematical model to predict the risk of human diseases spread between countries. Validation and application to the 2014 Ebola Virus Disease epidemic. *Bulletin of Mathematical Biology*, 77(9):1668-1704, 2015. doi:<https://doi.org/10.1007/s11538-015-0100-x>.
- [36] A.B. Kubik. *Parameter estimation for a mathematical model predicting the COVID-19 spread in the Àrea Metropolitana de Barcelona*. Ms Thesis, 2020. URL: <https://upcommons.upc.edu/handle/2117/328166>.
- [37] A. J. Kucharski, T. W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, R. M. Eggo et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *The Lancet Infectious Diseases*, 2020. doi:[https://doi.org/10.1016/S1473-3099\(20\)30144-4](https://doi.org/10.1016/S1473-3099(20)30144-4).

- [38] T. Liu, J. Hu, M. Kang, L. Lin, H. Zhong, J. Xiao et al. Transmission dynamics of 2019 novel coronavirus (2019-ncov). *bioRxiv*, 2020. doi:<https://doi.org/10.1101/2020.01.25.919787>.
- [39] R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, and J. Shaman. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science*, 2020. doi:<https://doi.org/10.1126/science.abb3221>.
- [40] W. Luo, M. Majumder, and Liu D. The role of absolute humidity on transmission rates of the COVID-19 outbreak. *MedRxiv*, 2020. doi:<https://doi.org/10.1101/2020.02.12.20022467>.
- [41] Ministerio de Ciencia e Innovación. Informe del GTM sobre el impacto de la COVID-19 en las personas mayores, con especial énfasis en las que viven en residencias. URL: https://www.ciencia.gob.es/stfls/MICINN/Ministerio/FICHEROS/Informe_residencias_GDT_MinisterioCyI.pdf, July 2020.
- [42] Ministero della Salute. Covid-19 - Situazione in Italia. URL: <http://www.salute.gov.it/portale/nuovocoronavirus/dettaglioContenutiNuovoCoronavirus.jsp?lingua=italiano&id=5351&area=nuovoCoronavirus&menu=vuoto>, 2020.
- [43] Ministero della Salute. Fase 2, nuovo decreto sulle riaperture: ecco cosa prevede. URL: <http://www.salute.gov.it/portale/nuovocoronavirus/dettaglioNotizieNuovoCoronavirus.jsp?lingua=italiano&menu=notizie&p=dalministero&id=4763>, May 2020.
- [44] Ministero della Salute. Nuovo coronavirus, situazione Italia. URL: <http://www.salute.gov.it/nuovocoronavirus>, 2020.
- [45] K. Y. Ng and M. M. Gui. COVID-19: Development of a robust mathematical model and simulation package with consideration for ageing population and time delay for control action and resusceptibility. *Physica D: Nonlinear Phenomena*, 411:132599, 2020. doi:<https://doi.org/10.1016/j.physd.2020.132599>.
- [46] C. N. Ngonghala, E. Iboi, S. Eikenberry, M. Scotch et al. Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel Coronavirus. *Mathematical Biosciences*, 325:108364, 2020. doi:<https://doi.org/10.1016/j.mbs.2020.108364>.
- [47] L. Ni, F. Ye, M. L. Cheng, Y. Feng et al. Detection of sars-cov-2-specific humoral and cellular immunity in COVID-19 convalescent individuals. *Immunity*, 52(6):971–977.e3, 2020. doi:<https://doi.org/10.1016/j.immuni.2020.04.023>.
- [48] L. Pérez, L. Pérez, M. Herranz et al. A complete analysis of the epidemiological scenario around a SARS-CoV-2 reinfection: previous infection events and subsequent transmission. Preprint (Version 1) available at *Research Square*, 13 November 2020. doi:<https://doi.org/10.21203/rs.3.rs-106167/v1>.
- [49] A. Remuzzi and G. Remuzzi. COVID-19 and Italy: what next? *Health Policy*, 395(10231):1225–1228, 2020. doi:[https://doi.org/10.1016/S0140-6736\(20\)30627-9](https://doi.org/10.1016/S0140-6736(20)30627-9).
- [50] H. I. Rhys. *Machine Learning with R, the tidyverse, and mlr*. Manning Publications, 2020. URL: <https://books.google.es/books?id=jRzYDwAAQBAJ>.
- [51] K. Roosa et al. Short-term Forecasts of the COVID-19 Epidemic in Guangdong and Zhejiang, China: February 13-23, 2020. *Journal of Clinical Medicine*, 9(2):596, 2020. doi:<https://doi.org/10.3390/jcm9020596>.
- [52] A. B. Ruiz, R. Saborido, and M. Luque. A preference-based evolutionary algorithm for multiobjective optimization: the weighting achievement scalarizing function genetic algorithm. *Journal of Global Optimization*, 62(1):101–129, 2015. doi:<https://doi.org/10.1007/s10898-014-0214-y>.

- [53] T. W. Russell, J. Hellewell, C. I. Jarvis, K. van Zandvoort, S. Abbott et al. Estimating the infection and case fatality ratio for COVID-19 using age-adjusted data from the outbreak on the Diamond Princess cruise ship. *medRxiv*, 2020. URL: <https://www.medrxiv.org/content/early/2020/03/09/2020.03.05.20031773>, doi:10.1101/2020.03.05.20031773.
- [54] P. Van den Driessche and J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences*, 180(1-2):29–48, 2002. doi:<https://doi.org/10.1101/2020.02.22.20025791>.
- [55] M. A. C. Vollmer, S Mishra, H. J .T. Unwin, A. Gandy et al. Using mobility to estimate the transmission intensity of COVID-19 in italy: a subnational analysis with future scenarios. *Imperial College London*, 2020. doi:<https://doi.org/10.25561/78677>.
- [56] M. Wang, M. D. Aili Jiang, and L. Gong. Temperature Significantly Change COVID-19 transmission in 429 cities. *MedRxiv*, 2020. doi:<https://doi.org/10.1101/2020.02.22.20025791>.
- [57] L. Wang, Y. Wang, Y. Chen, and Q. Qin. Unique epidemiological and clinical features of the emerging 2019 novel coronavirus pneumonia (COVID-19) implicate special control measures. *Journal of Medical Virology*, 2020. doi:<https://doi.org/10.1002/jmv.25748>.
- [58] J. Wang, M. Zhou, and F. Liu. Reasons for healthcare workers becoming infected with novel coronavirus disease 2019 (COVID-19) in China. *The Journal of hospital infection*, 105(1):100–101, 2020. doi:<https://doi.org/10.1016/j.jhin.2020.03.002>.
- [59] World Bank. World Bank Open Data. URL: <https://data.worldbank.org/>, March 2020.
- [60] World Health Organization. Coronavirus disease (COVID-2019) Situation reports. URL: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>, March 2020.
- [61] World Health Organization. Coronavirus disease (COVID-2019) Situation reports 73. URL: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200402-sitrep-73-covid-19.pdf?sfvrsn=5ae25bc7_6, April 2020.
- [62] World Health Organization. Coronavirus disease (COVID-2019) Situation reports. URL: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>, June 2020.
- [63] World Health Organization. Director-general’s opening remarks at the media briefing on COVID-19 - 11 march 2020. URL: <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>, March 2020.
- [64] World Health Organization. Report of the WHO-China Joint Mission on Coronavirus Disease 2019. URL: <https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>, February 2020.