



Compartmentalized mathematical model to predict future number of active cases and deaths of COVID-19

Osmar Pinto Neto^{1,2,3} · José Clark Reis² · Ana Carolina Brisola Brizzi^{1,2} · Gustavo José Zambrano² · Joabe Marcos de Souza^{2,4} · Wellington Pedroso^{1,2} · Rodrigo Cunha de Mello Pedreiro^{1,5,6} · Bruno de Matos Brizzi² · Ellysson Oliveira Abinader⁷ · Renato Amaro Zângaro^{1,3}

Received: 21 April 2020 / Accepted: 20 August 2020 / Published online: 30 August 2020

© Sociedade Brasileira de Engenharia Biomedica 2020

Abstract

Introduction In December 2019, China reported a series of atypical pneumonia cases caused by a new Coronavirus, called COVID-19. In response to the rapid global dissemination of the virus, on the 11th of Mars, the World Health Organization (WHO) has declared the outbreak a pandemic. Considering this situation, this paper intends to analyze and improve the current SEIR models to better represent the behavior of the COVID-19 and accurately predict the outcome of the pandemic in each social, economic, and political scenario.

Methodology We present a generalized Susceptible-Exposed-Infected-Recovered (SEIR) compartmental model and test it using a global optimization algorithm with data collected from the WHO.

Results The main results were: (a) Our model was able to accurately fit the either deaths or active cases data of all tested countries using optimized coefficient values in agreement with recent reports; (b) when trying to fit both sets of data at the same time, fit was good for most countries, but not all. (c) Using our model, large ranges for each input, and optimization we predict death values for 15, 30, 45, and 60 days ahead with errors in the order of 5, 10, 20, and 80%, respectively; (d) sudden changes in the number of active cases cannot be predicted by the model unless data from outside sources are used.

Conclusion The results suggest that the presented model may be used to predict 15 days ahead values of total deaths with errors in the order of 5%. These errors may be minimized if social distance data are inputted into the model.

Keywords COVID-19 · Compartmental model · Active cases · Deaths · Epidemiological model predictions

Introduction

In December 2019, in China, a series of atypical pneumonia cases have emerged caused by a new Coronavirus, nowadays officially called COVID-19 by the World Health Organization (WHO). It has spread rapidly throughout the country, having had its epicenter the city of Wuhan, in which 82,249 people were infected and 3341 people have died. In response to the rapid global dissemination of the virus, on the 11th of Mars the WHO has declared the outbreak a pandemic. Since then, the global impact of the COVID-19 became a great threat to the public health. Considering this emergency, different areas of science need to focus its attention to the challenges imposed by this new Coronavirus. In such scenarios, it is imperative the necessity of new, improved, and specific mathematical modeling.

There are many uncertainties regarding the gravity of the infection caused by COVID-19. Nevertheless, based on

✉ Osmar Pinto Neto
osmarpintoneto@hotmail.com

¹ Biomedical Engineering Department, Anhembi Morumbi University, Sao Paulo, SP, Brazil

² Arena235 Research Lab – São José dos Campos, Sao Jose dos Campos, SP, Brazil

³ Center for Innovation, Technology and Education – CITE, Parque Tecnológico de São José dos Campos, Estrada Dr. Altino Bondensan, 500, Sao Jose dos Campos, SP 12247-016, Brazil

⁴ Departamento de Engenharia Aeronáutica, Universidade de São Paulo, Sao Paulo, SP, Brazil

⁵ Estácio de Sá University, Nova Friburgo, RJ, Brazil

⁶ Santo Antônio de Pádua College, Santo Antonio de Padua, RJ, Brazil

⁷ Instituto Abinader, Manaus, AM, Brazil

epidemiological investigations, the period of incubation is 1 to 14 days, more evidently between 3 and 7 days, and the virus is contagious still on its latency period (Guo et al. 2020). The majority of the infected adults and children have developed mild symptoms alike those of a common cold, and some patients evolved rapidly to an acute respiratory discomfort, followed by respiratory failure, multiple organ failure, and death. The probability of death in the USA, according to the Centers for Disease Control and Prevention (CDC), has ranged from 0.5% to ages of 45–54 to 1.4% to ages 55–64 and growing on a constant rate with age. Pre-existing comorbidities, affecting the vulnerability to the infection, also increase the probability of death.

COVID-19 seems to have a relatively higher rate of transmissibility when compared with other Coronavirus infections and, to better understand it, it is very important to multiple factors (Chen et al. 2020; Wang et al. 2020). Family environment, age, and wealth distribution are essential factors related to the transmission and mortality rate of COVID-19 (Walker et al. 2020). Another important that must be considered, especially in the mortality rate, is the number of hospital beds and the capacity of Intensive Care Units (ICU). The relationship between the age ranges that require attention and the mortality rate by infection was scrutinized by China and, assuming that 30% of the hospitalized will demand intensive care (ICU), and among those 50% will die, it has been calculated the demand for hospital beds assuming the average stay of 16 days in the hospital (Ferguson et al. 2020). The experience of COVID-19 in many countries, concerning medical assistance, has indicated that the demand for hospital beds and the need for mechanical ventilation have overcome the availability of those in countries with higher per capita income. Therefore, the consequences in countries with scarcity of these services are expected to be larger (Walker et al. 2020).

Previous experiences in some countries highlight the need to anticipate the impacts of the pandemic outbreak and to develop researches with epidemiological models. These mathematical models are necessary to the comprehension of the present outbreak's behavior, so that countries might develop strategies to minimize the impacts on the healthcare system and preserve life (Wu et al. 2020; Peng et al. 2020). As an example, public administrators may find comprehensive ground to define policies such as enforcing social distance measures, the available versus need of laboratorial tests, planning for hospital beds, and health system resources.

In the absence of a vaccine, mathematical modeling may assess the effectiveness of non-pharmaceutical interventions and in its role in decreasing population contact and viral transmission to control the pandemic outbreak. China has managed to control the outbreak with the deployment of isolation of its cases and social distancing of the population (Ferguson et al. 2020).

There are some non-pharmaceutical strategies to control an outbreak, such as containment, mitigation, and suppression.

When the containment measures fail to control the outbreak, mitigation and suppression strategies may be adopted to postpone and mitigate its effects on society and the healthcare system. Mitigation will concentrate in retarding but not necessarily impeding the spread of the outbreak, reducing the peak of medical assistance and protecting the higher risk groups. Suppression aims to reverse the outbreaks' growth, diminishing the number of cases and maintaining this frame for an indefinite time, through more extreme measures, such as quarantining, police enforcement, mass testing, compulsory notification, and finance support to the population in isolation, among other actions. (Ferguson et al. 2020; Walker et al. 2020).

Concerning mathematical modeling, which supplies detailed mechanisms about the outbreak dynamics, the Susceptible-Exposed-Infected-Recovered (SEIR) epidemiological model is widely adopted to characterize a pandemic caused by COVID-19. For instance, this method was used for decision making in Hubei, Wuhan, and Beijing (Peng et al. 2020).

This paper intends to analyze and improve up on the traditional SEIR model by adding important new compartments, as well as consider the effects of non-pharmaceutical interventions and the possibility of death coming from lack of available ICU beds. Moreover, such a model can also be used to prototype and analyze the cause/effect relation of a multitude of actions and public health strategies, so the most effective ones can be chosen for each country, city, or province. Since every affected region is different, it is of utmost importance to help organizations to determine not only the number of active cases but also the number of hospital beds, and ICUs will be needed at a certain point in time to maximize the usage of public resources.

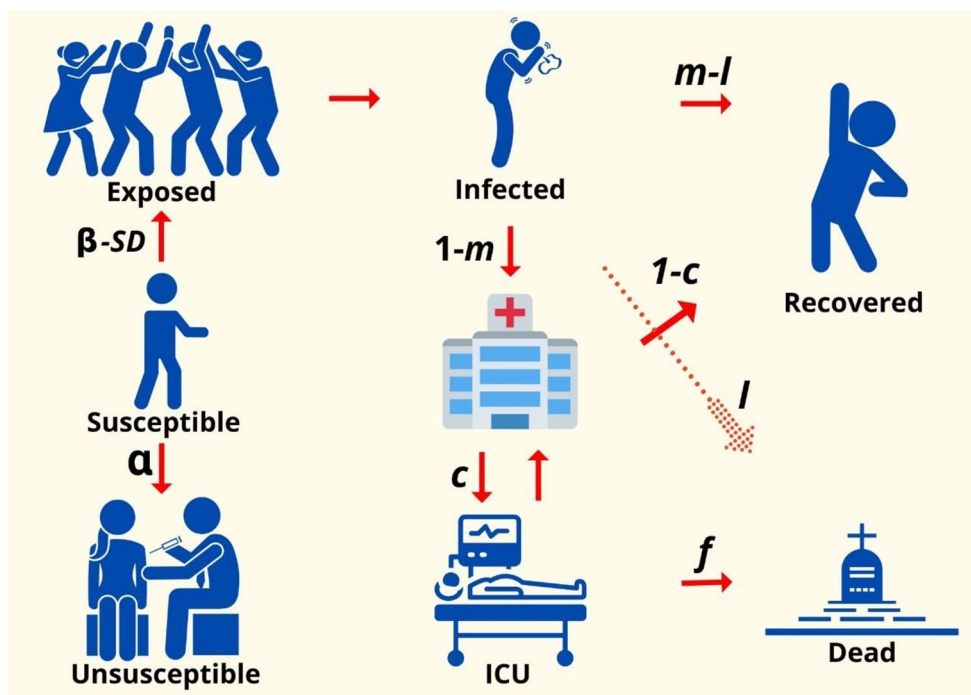
Methodology

The SUEIHCDR model

We present a generalized SEIR compartmental model using novel and recently suggested ideas and concepts (APMonitor Optimization Suite 2020; Peng et al. 2020, University of Basel 2020). An application of our model using real mobility data to investigate different future projections for the USA has been recently reported (Kennedy et al. 2020). It is composed of eight compartments: Susceptible, Unsusceptible, Exposed, Infected, Hospitalized, Critical, Dead, and Recovered (SUEIHCDR; Fig. 1).

The model assumes, at first that, the whole population is susceptible (Eq. 1) to the disease. As time progresses, a susceptible person can either become exposed (Eq. 5) to the virus or unsusceptible (Eq. 2).

Fig. 1 SUEIHCDR model info graphic description; it is composed of eight compartments Susceptible, Unsusceptible, Exposed, Infected, Hospitalized, Critical, Dead, and Recovered. β is the infection rate, SD is a social distancing factor, α is a protection rate, m is the fraction of infectious that are asymptomatic, $1 - m$ is the percentage of the infected go hospitalized, l is the percentage of infected people that may die without hospitalization, $1 - c$ is the percentage of hospitalized people that recovers, c is the fraction of hospitalized that becomes critical cases needing to go an intensive care unit (ICU), and f is the fraction of people in critical state that dies



$$\frac{dS(t)}{dt} = -\frac{(1-SD(t))\beta S(t)I(t)}{N_{pop}} - \alpha(t)S(t) \tag{1}$$

$$\frac{dU(t)}{dt} = \alpha(t)S(t) \tag{2}$$

where $I(t)$ is the number of infectious people at time t , N_{pop} is the population of the country, β is the infection rate, α is a protection rate, and SD is a social distancing factor.

As in Peng et al. (2020) we introduced a protection rate α factor to our susceptible equation (Eq. 1). This protection rate was introduced to account for possible decreases in the number of susceptible people to the virus caused by factors other than social distancing, such as the usage of face masks, better hygiene, more effective contact tracing, and possible vaccines and or drugs that may prevent infection. Different from the aforementioned study of Peng et al. (2020), however, we varied α across time (Eq. 3). This time variation was introduced to reliably model people’s behavior, who are not commonly too concerned about the disease in the earlier stages of the epidemic, but as the number of infected and deaths increases, become more cautious about the virus.

$$\alpha(t) = \alpha_0 \frac{\log(t + 1)}{\log(t_f)} \tag{3}$$

where α_0 is the reference value that is the maximum value and t_f is the final time for the prediction.

Furthermore, we also introduced a social distancing factor SD , which also varies with time (Eq. 4). Social distancing was modeled as a logistic curve so that the model could account for

the date (t_{sd}) when a possible quarantine measurement starts. As mentioned before, real data for mobility can be used in our model when available; using real mobility data have been shown to be important when long-term future projections are intended (Kennedy et al. 2020).

$$SD(t) = SD_0 \frac{1}{1 + e^{-(t-t_{sd})}} \tag{4}$$

where SD_0 is SD reference value, that is the maximum value, and t_{sd} is the time the SD increases until reaching SD_0 .

Exposed people become infectious after an incubation time of $1/\gamma$ (Eq. 5).

$$\frac{dE(t)}{dt} = \frac{(1-SD(t))\beta S(t)I(t)}{N_{pop}} - \gamma E(t) \tag{5}$$

Infected people stay infected for a period of $1/\delta$ (Eq. 6) days and can have three different outcomes. Considering m as a specific parameter to account the fraction of infectious that is asymptomatic, it is possible to determine that a percentage of the infected ($1 - m$) go hospitalized, another percentage of them (l) may die without hospitalization, and the rest of them ($m - l$) recover. l was introduced as a function of time (Eq. 7) so that the time when hospital bed became unavailable could be modeled (t_m), as well as the duration that hospital was full (dur).

$$\frac{dI(t)}{dt} = \gamma E(t) - \delta I(t) \tag{6}$$

Table 1 Input coefficients to the model and respective ranges

Coeff.	Lower bound	Higher bound
α	0.01	0.12
β	0.5	1.2
γ	0.5	5.00
δ	0.07	0.50
ζ	0.20	0.33
ε	0.05	0.14
m	0.65	0.85
l	0.005	0.02
rl	0.00	0.16
t_l	0.00	40
dur	5	40
c	0.10	0.50
f	0.35	0.55
$E0$	$E0/2$	$2E0$
$I0$	$I0/2$	$2I0$
$H0$	$H0/2$	$2H0$
$C0$	$C0/2$	$2C0$
$Rec0$	$Rec0/2$	$2Rec0$
$D0$	$D0/2$	$2E0$
SD	0.00	0.75
m_{sd}	0.00	40

$$l(t) = \begin{cases} l_0 \frac{1}{1 + e^{-rl(t-t_l)}}, & t < 2t_l + dur \\ 0.95^{l_0(t-2t_l+dur)}, & t \geq 2t_l + dur \end{cases} \quad (7)$$

where l_0 is the inclination of the angular coefficient of the ramp up until reaching the maximum value reference value and t_l is the time when people started dying due the lack of available ICUs.

Hospitalized people (Eq. 8) stay hospitalized for $1/\zeta$ days and can either recover ($1 - c$) or become critical (c —specific parameter to account the fraction of hospitalized that becomes critical cases) needing to go an intensive care unit (ICU).

$$\frac{dH(t)}{dt} = (1-m)\delta I(t) + (1-f)\varepsilon C(t) - \zeta H(t) \quad (8)$$

where ε is the inverse of the time people stay in the ICU.

A person stays on average $1/\varepsilon$ in the ICU (Eq. 9) and can either go back to the hospital ($1 - f$) or die (f —specific parameter to account the fraction of people in critical state that died).

$$\frac{dC(t)}{dt} = c\zeta H(t) - \varepsilon C(t) \quad (9)$$

Therefore, recovered people (Eq. 10) can either come straight from infection when the case is mild ($m - l$) or from the hospital when the case is no critical ($1 - c$).

Table 2 Optimized inputs for Germany, Brazil, Spain, Italy, South Korea, Portugal, Switzerland, Thailand, and USA, respectively, for June 18, 2020. μ stands for the mean across countries and STD for the standard deviation

Cf	BR	GER	ITA	POR	KOR	SPA	SWI	THA	USA	μ	SD
α	0.016	0.034	0.024	0.022	0.036	0.023	0.034	0.034	0.020	0.027	0.007
β	0.63	0.61	0.63	0.62	0.60	0.57	0.57	0.40	0.66	0.59	0.07
γ	1.02	0.76	1.25	1.23	0.70	0.80	1.18	0.93	1.12	1.00	0.20
δ	0.09	0.11	0.10	0.11	0.13	0.11	0.13	0.14	0.13	0.12	0.02
ζ	0.26	0.27	0.16	0.25	0.25	0.22	0.25	0.26	0.23	0.24	0.03
ε	0.09	0.11	0.09	0.08	0.10	0.06	0.09	0.10	0.08	0.09	0.01
m	0.90	0.90	0.91	0.93	0.96	0.85	0.95	0.96	0.89	0.92	0.04
l	0.004	0.010	0.006	0.006	0.003	0.008	0.010	0.002	0.003	0.006	0.003
rl	0.08	0.08	0.08	0.08	0.11	0.07	0.09	0.09	0.06	0.08	0.01
t_l	17.47	24.80	23.71	17.95	25.35	22.22	19.79	16.75	18.60	20.74	3.14
dur	21.65	24.38	26.10	23.25	19.84	25.39	22.49	24.33	24.45	23.54	1.84
c	0.29	0.26	0.25	0.19	0.21	0.27	0.33	0.17	0.25	0.25	0.05
f	0.31	0.35	0.32	0.33	0.32	0.28	0.42	0.21	0.31	0.32	0.05
$E0$	1357	1059	1967	922	1036	13,223	1134	487	736	2436	3834
$I0$	987	1289	775	1084	680	15,784	829	450	496	2486	4709
$H0$	1038	1386	814	1138	626	16,661	843	448	546	2611	4975
$C0$	31	39	36	20	37	600	31	38	12	94	179
$Rec0$	0	0	0	0	0	0	0	0	0	0	0
$D0$	0	0	0	0	0	833	0	0	0	93	262
SD	0.60	0.53	0.57	0.62	0.64	0.60	0.56	0.51	0.51	0.57	0.04
m_{sd}	5	9	7	3	5	3	7	6	17	7	4
F	1.8	4.9	3.5	6.1	1.4	1.9	2.4	2.1	2.0	2.9	1.5

$$\frac{dR(t)}{dt} = (m-l)\delta I(t) + (1-c)\zeta H(t) \tag{10}$$

Death (Eq. 11) arises either from lack of available treatment (l), or from critical cases in the ICU (f).

$$\frac{dD(t)}{dt} = l\delta I(t) + f\varepsilon C(t) \tag{11}$$

At last, the effective reproduction number R_t (Eq. 12) of our model can be estimated as

$$R_t(t) = \frac{(1-SD(t))\beta}{\delta} \left(1 - \frac{\int_0^t \alpha(t)S(t)}{N_{pop}} \right) \tag{12}$$

Solving and testing the model

We used the fourth order Runge-Kutta numerical method to solve our system of ordinary differential equation in MATLAB (Mathworks Inc.R17a). To test our model we gathered active cases, recovered cases, accumulated deaths, and tests per million people data from the WHO of ten different countries in different stages of the epidemic: Germany, Brazil,

Spain, Italy, South Korea, Portugal, Switzerland, Thailand, and USA. Lack of testing and under-notification of active cases has been largely reported for the COVID-19 (Hasell et al. 2020;WorldoMeter, 2020; UFPel, 2020); in consequence, active cases data were corrected by a factor. The correction factor was found via optimization, as described in the next paragraph, using a range of possibilities estimated based on previous reports. Lower bound was determined considering the death rate of the country to be as described in Verity et al. (2020), corrected by age (young: 0.32%; older adults (60+): 6.4%; senior older adults (80+): 13.4%). Upper bound was set by considering the same age proportional differences in death rate, as previously mentioned, but adjusting death rate of each country by the death rate in Iceland (country with the greatest percentage of test per inhabitant Gudbjartsson et al. 2020).

We used a custom build MATLAB global optimization algorithm using Monte Carlo iterations algorithm and multiple local minima searches. The algorithm was tested for the best solution considering 21 different inputs to the model within ranges obtained from the WHO and several publications (Liu et al. 2020; Ranjan 2020; Wu et al. 2020; Table 1) and 1 correction factor (F) for the active cases (Table 2). The algorithm was used to minimize a goal function (J) as a combination of Active Cases and Death time series (Eq. 13).

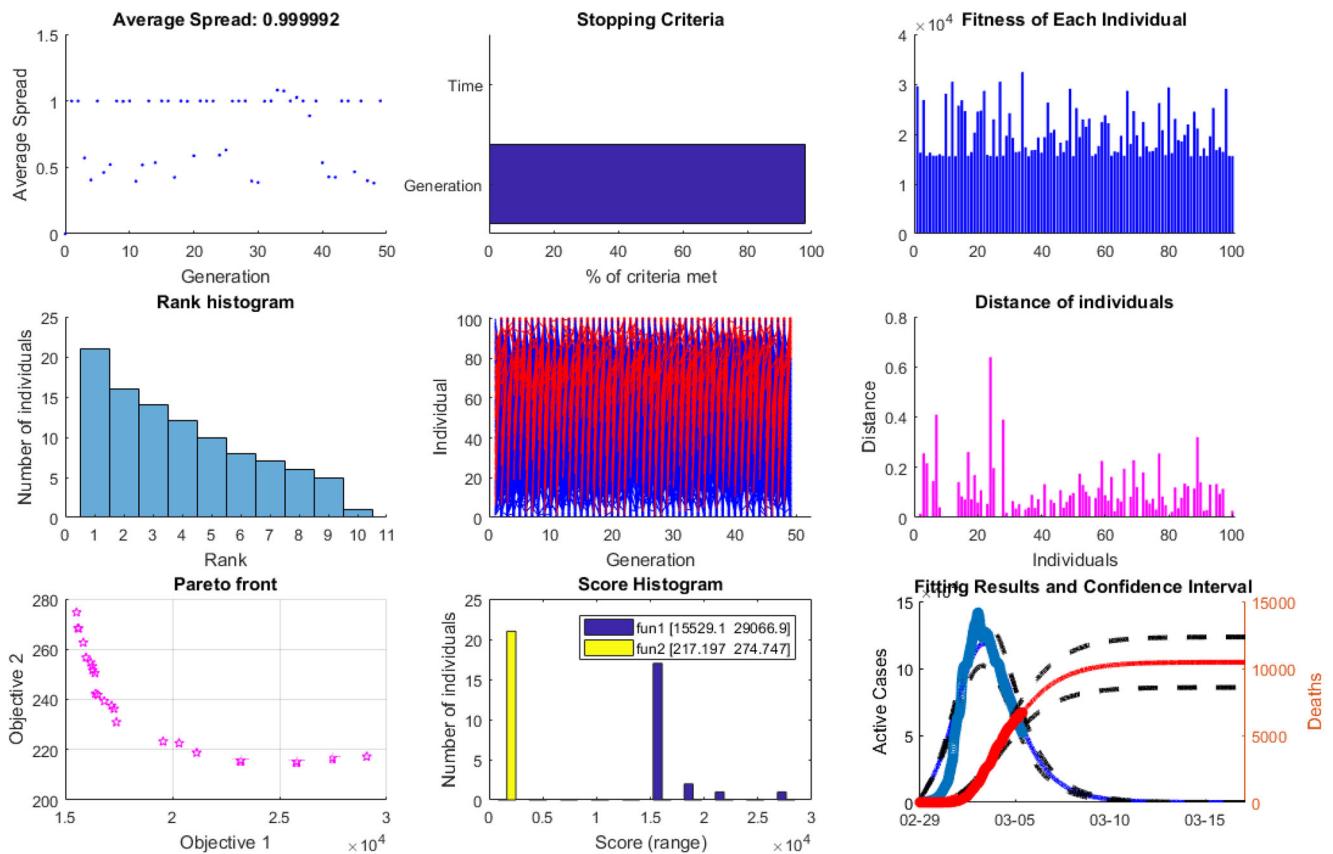


Fig. 2 Optimization algorithm results after 10,000 runs. The Pareto front was determined considering two simultaneous objectives active cases RMSE and deaths RMSE. The best 100 solutions were used to initialize the MATLAB multi-objective genetic algorithm

$$\begin{cases} J = (1-p)\text{RMSE}(\text{Active Cases}) + p\text{RMSE}(\text{Deaths}) \\ p = \frac{\text{rms}(\text{Active Cases})/\text{rms}(\text{Death})}{\text{rms}(\text{Active Cases})/\text{rms}(\text{Death}) + 1} \end{cases} \quad (13)$$

Data under 500 active cases were discarded. Initial values for each compartmental parameter had ranges proportional to the following initial values (Table 1): infected initial values (I_0) were determined as the corrected actives cases first value greater than 50; exposed initial values (E_0) were $0.5 \times I_0$; hospitalized initial values (H_0) were $0.2 \times I_0$; critical cases initial values (C_0) were $0.7 \times I_0$; death initial values (D_0) were obtained from the accumulated deaths real data; similarly, recovered initial values (Rec_0) were obtained from the recovered real data. Optimization algorithm results were considered after 10,000 runs; from them, the 100 best solutions were used as initial population for a multi-objective genetic algorithm (MATLAB function: *gamultiobj*) to determine a Pareto front of solutions considering two simultaneous objectives RMSE (active cases) and RMSE (Deaths) (Fig. 2). Lower and upper bounds for the genetic algorithm were set as 40% variations to

the best solution found out of the previous optimization algorithm runs, and 100 generations were created.

All fitting processes were done for data from the day 1 of the outbreak for each country up to April 18, May 3, May 18, June 3, and June 18, 2020, to test the accuracy of the future predictions that can be made based on the model and optimization results. Furthermore, 2% perturbations to the model coefficients were used to determine via Monte Carlo a 95% confidence interval for the results. Results are presented as mean (standard deviation).

Results

Besides introducing more compartments than a traditional *SEIR* model (e.g., hospitalized) the three main differences of our model *SUEIHCDR* to a standard *SEIR* model is the addition of α *SD* and *l*. Our results suggest that our model was able to accurately fit the data of all countries when one goal is considered (Figs. 3 and 4). However, when we tried to fit the model two both accumulated deaths and active cases, we found that we could not reproduce with the same accuracy the data for all analyzed countries (Fig. 5).

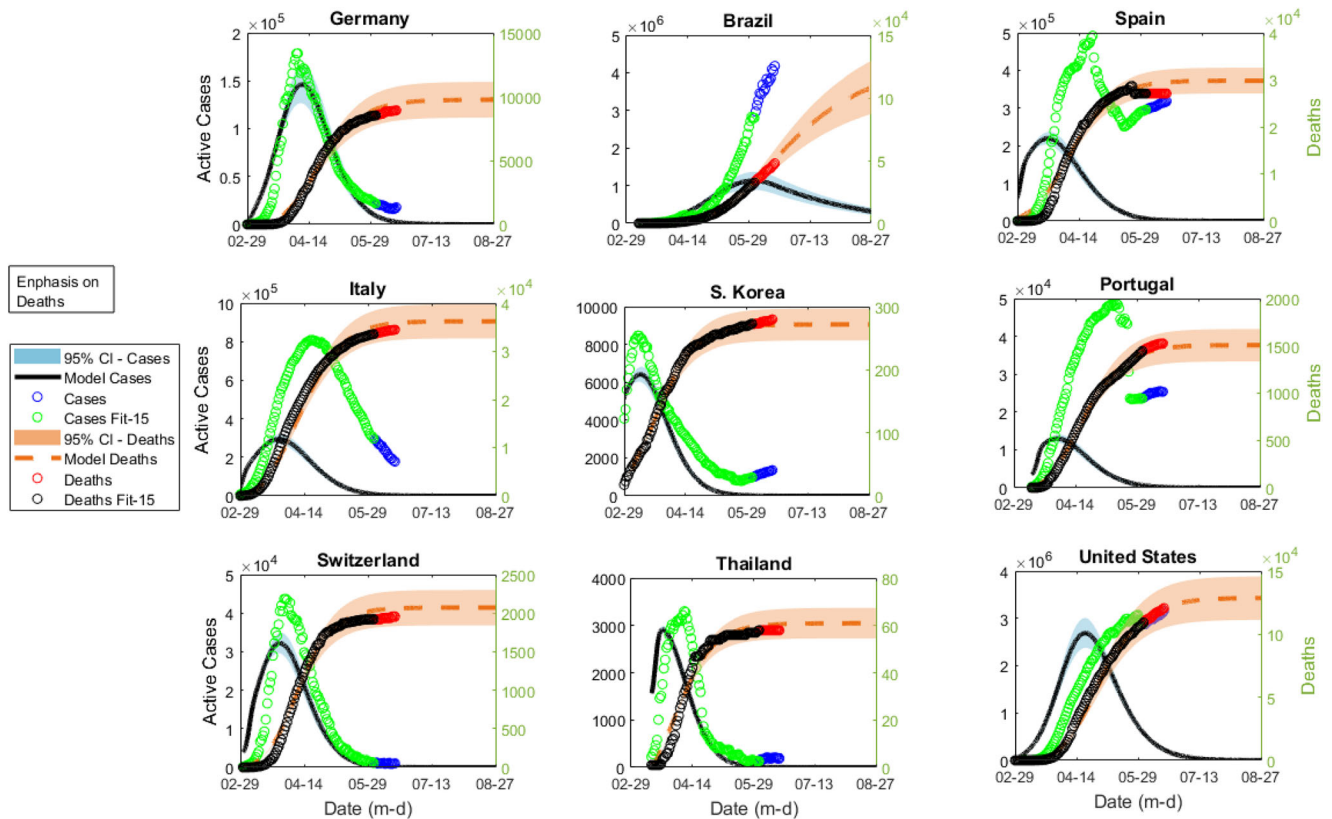


Fig. 3 Model results for active cases and accumulated deaths for all studies countries, considering minimizing deaths RMSE. Optimization was done considering end-date June 3 (black circles for deaths and green

circle for active cases). The red circles (deaths) and blue circles (active cases) indicate real data up to June 18

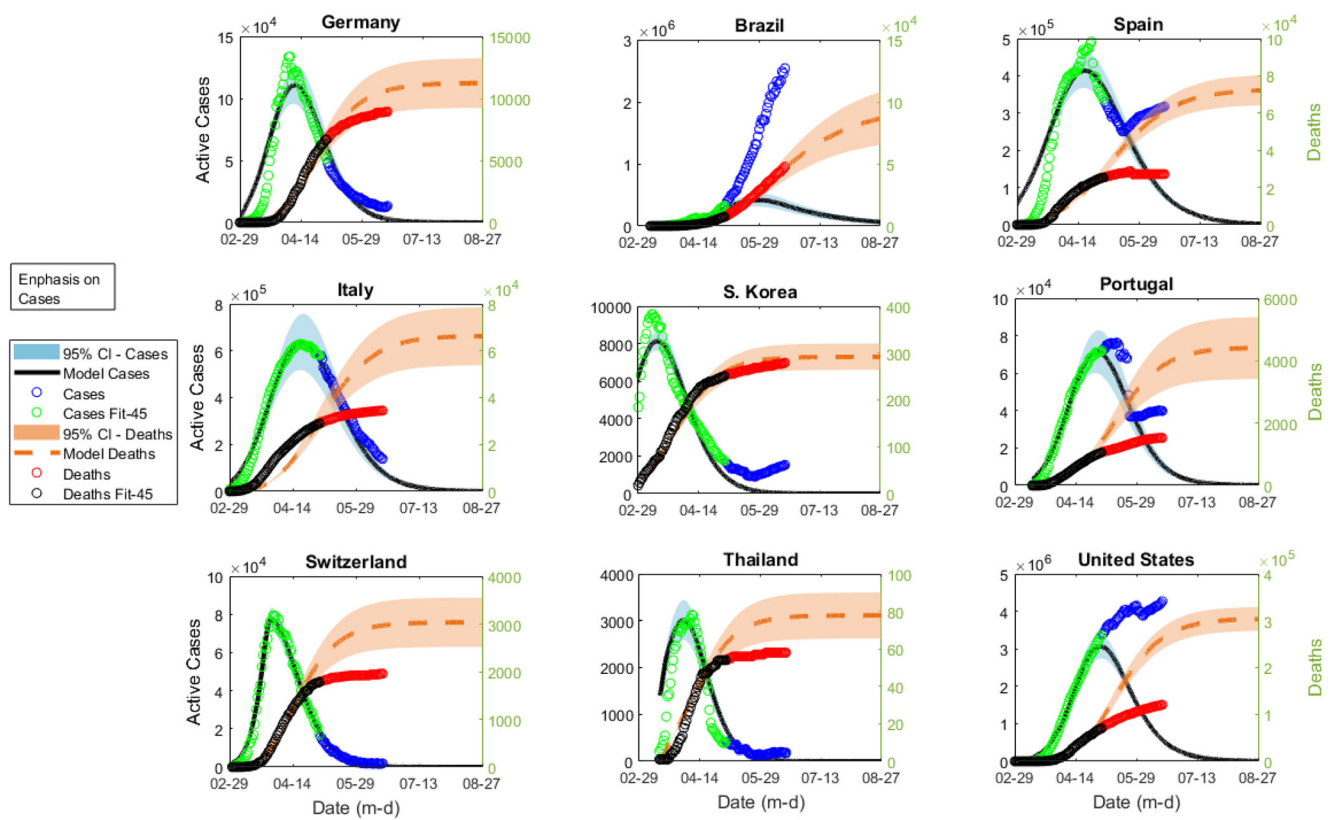


Fig. 4 Model results for active cases and accumulated deaths for all studies countries, considering minimizing active cases RMSE. Optimization was done considering end-date May 3 (black circles for

deaths and green circle for active cases). The red circles (deaths) and blue circles (active cases) indicate real data up to June 18

Table 2 shows the optimization parameter results for June 18, 2020, for Germany, Brazil, Spain, Italy, South Korea, Portugal, Switzerland, Thailand, and the USA considering the solution from the Pareto front (Fig. 2) that minimized J (Eq. 12). Mean protective rate (α) was 0.027 (0.007); mean infectious rate (β) was 0.59 (0.07); mean fraction of infectious that are asymptomatic or mild (m) was 0.92 (0.04); the mean fraction of infectious people that died with no treatment (l) was 0.006 (0.003); the fraction of severe cases that turn critical (c) was 0.25 (0.05); the mean fraction of critical cases that are fatal (f) was 0.32 (0.05), and the mean social distancing parameter (SD) was 0.57 (0.04). Table 3 shows the inverse values of γ , δ , ζ , and ε : the mean latent period was 1 (0.2) days; the mean infectious period was 8.9 (1.5); the mean hospitalized period was 4.3 (0.7) days, and the mean period in ICU was 11.5 (2.1) days. The basic reproduction number (R_0) was 2.24 (0.52), and the death rate was 1.2 (0.5)%.

Figures 3, 4, and 5 show the model results for all studied countries. Figure 3 shows results considering the two-goal optimization Pareto front that minimized death RMSE. Optimization was done considering end-date June 3 (black circles for deaths and green circle for active cases). The red circles (deaths) and blue circles (active cases) indicate “future” real data for the next 15 days. Similarly, Fig. 4 shows the

results for the two-goal optimization, but now minimizing active cases RMSE and end-date May 3. Finally, Fig. 5 shows the results considering the solution from Pareto front that minimized J for end-date May 18th.

Table 4 shows the model future projection of 15, 30, 45, and 60 days for total number of infected, deaths, hospitalized, peak hospitalization, ICU patients, peak day ICU, and recovered patients. Results indicate deaths in the thousands for every country but Korea and Thailand. USA has a peak day of more than 100 thousand hospitalized patients. Additionally, Spain projects almost 2 million recovered people by the end of August 2020. According to the model estimations, Brazil will have more than 200 thousand ICU patients treated by the end of July 18 and 70 thousand more in the following 30 days; peak day will demand 35 thousand ICU beds.

Finally, Tables 5, 6, and 7 shows the percentage errors comparing model results to real data for the day of the analyses and future projections of 15, 30, 45, and 60 days. Table 5 shows the results for the optimization minimizing J , Table 6 shows the minimizing death RMSE, and Table 7 shows the minimizing active case RMSE. Thirty-day projections were performed twice: first, as for the other time windows, considering future date June 18, and second considering future data May 18. As expected, errors got larger for farther into the

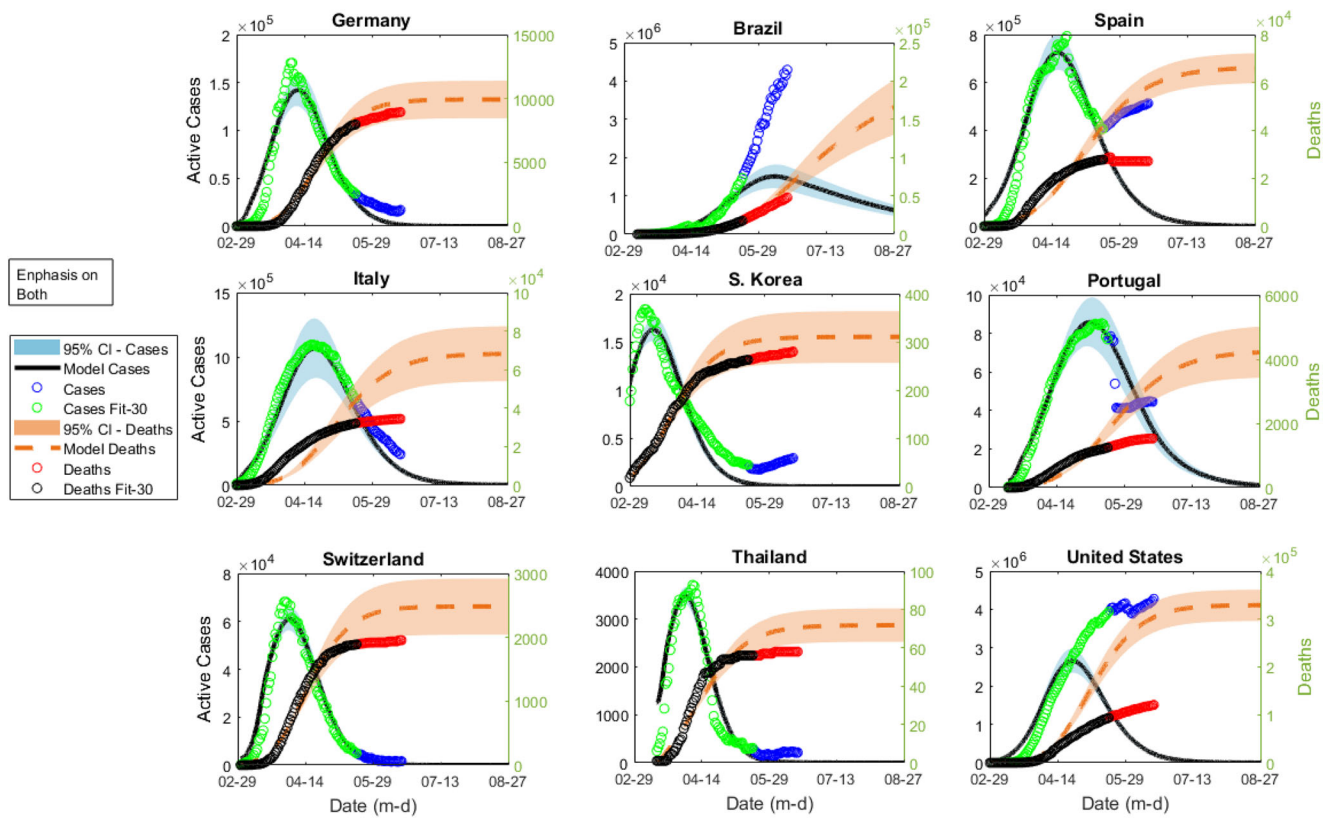


Fig. 5 Model results for active cases and accumulated deaths for all studies countries, considering minimizing both the active cases and the death RMSE. Optimization was done considering end-date May 18

(black circles for deaths and green circle for active cases). The red circles (deaths) and blue circles (active cases) indicate real data up to June 18

future projections. In general, projected deaths had smaller percentage errors. Because of recent re-opening of countries such as Portugal and Spain 30-day future projections considering data from 60 days ago to estimate 30 days ago yielded

better results than data from 30 days ago to estimate present day (May 18). As it can be seen in Fig. 4, there was a sudden increase in the number of active cases for both countries in the past days that were not predicted by the model whose active case curves kept on a steady decline.

Table 3 Inverse of the model optimized coefficients of γ , δ , ζ , and ε representing latent, infectious, hospitalization, and critical cases mean duration in days, as well as the model estimated basic reproductive number (R_0) and the death rate (DR) for June 18, 2020, for Germany, Brazil, Spain, Italy, South Korea, Portugal, Switzerland, Thailand, and USA, respectively. μ stands for the mean across countries and STD for the standard deviation

Country	R_0	DR	Latent	Infectious	Hospitalized	Critical
BR	2.94	0.012	1.0	11.6	3.9	10.6
GER	2.50	0.017	1.3	8.9	3.7	8.9
ITA	2.87	0.013	0.8	10.5	6.2	11.1
POR	2.18	0.010	0.8	9.2	4.1	11.8
KOR	1.71	0.006	1.4	7.8	4.0	10.5
SPA	2.14	0.020	1.3	9.4	4.5	16.5
SWI	1.96	0.015	0.8	7.8	4.1	11.4
THA	1.38	0.005	1.1	7.0	3.9	10.0
USA	2.45	0.012	0.9	7.7	4.4	12.4
μ	2.24	0.012	1.0	8.9	4.3	11.5
STD	0.52	0.005	0.2	1.5	0.7	2.1

Discussion

Considering the rapid growing COVID-19 pandemic and the necessity of modeling the phenomenon to make future predictions in the number of cases, deaths, but ultimately in the number of hospital and ICU beds, we present a novel generalized SEIR compartmental model with the addition of the unsusceptible, hospitalized, critical, and dead compartments. Furthermore, we introduce three new parameters to the model (α , SD , and l). We tested our model using a global optimization algorithm and data collected from the WHO for several countries. Our main findings were as follows: (a) our model was able to accurately fit the either deaths or active cases data of all countries tested independent of what stage of the epidemic they were using optimized coefficient values in agreement with recent reports; (b) when trying to fit both sets of data at the same time, fit was good for some countries, but not for all; (c) using our model, large ranges for each input, and

Table 4 Model results for future projection (in thousands of people) of 15, 30, 45, and 60 days for total number of infected, deaths, hospitalized, peak hospitalization, ICU patients, peak day ICU, and recovered patients

Country	Days	Infect	Deaths	Hosp	peak_Hos	ICU	peak_ICU	REC
BR	15	7.21E + 06	7.33E + 04	6.61E + 05	4.65E + 04	1.55E + 05	3.46E + 04	6.05E + 06
GER	15	4.92E + 05	8.81E + 03	5.79E + 04	3.90E + 03	1.51E + 04	2.27E + 03	4.83E + 05
ITA	15	3.47E + 06	4.94E + 04	3.84E + 05	3.72E + 04	9.14E + 04	1.53E + 04	3.39E + 06
POR	15	2.65E + 05	2.66E + 03	2.19E + 04	1.34E + 03	4.01E + 03	6.95E + 02	2.58E + 05
KOR	15	4.29E + 04	2.82E + 02	2.78E + 03	6.26E + 02	6.17E + 02	1.01E + 02	4.26E + 04
SPA	15	1.86E + 06	3.88E + 04	3.69E + 05	2.50E + 04	9.50E + 04	2.13E + 04	1.81E + 06
SWI	15	1.40E + 05	2.22E + 03	8.89E + 03	8.43E + 02	2.97E + 03	5.40E + 02	1.38E + 05
THA	15	1.23E + 04	5.95E + 01	1.10E + 03	4.48E + 02	2.22E + 02	6.68E + 01	1.22E + 04
USA	15	1.01E + 07	1.32E + 05	1.38E + 06	1.07E + 05	3.37E + 05	6.85E + 04	9.94E + 06
BR	30	8.11E + 06	9.18E + 04	8.04E + 05	4.65E + 04	2.00E + 05	3.46E + 04	7.14E + 06
GER	30	4.92E + 05	8.85E + 03	5.81E + 04	3.90E + 03	1.52E + 04	2.27E + 03	4.83E + 05
ITA	30	3.48E + 06	5.07E + 04	3.91E + 05	3.72E + 04	9.50E + 04	1.53E + 04	3.41E + 06
POR	30	2.67E + 05	2.77E + 03	2.25E + 04	1.34E + 03	4.25E + 03	6.95E + 02	2.62E + 05
KOR	30	4.29E + 04	2.82E + 02	2.78E + 03	6.26E + 02	6.17E + 02	1.01E + 02	4.26E + 04
SPA	30	1.87E + 06	3.97E + 04	3.73E + 05	2.50E + 04	9.81E + 04	2.13E + 04	1.82E + 06
SWI	30	1.40E + 05	2.23E + 03	8.91E + 03	8.43E + 02	3.00E + 03	5.40E + 02	1.38E + 05
THA	30	1.23E + 04	5.96E + 01	1.10E + 03	4.48E + 02	2.22E + 02	6.68E + 01	1.22E + 04
USA	30	1.01E + 07	1.35E + 05	1.40E + 06	1.07E + 05	3.46E + 05	6.85E + 04	9.97E + 06
BR	45	8.82E + 06	1.07E + 05	9.20E + 05	4.65E + 04	2.39E + 05	3.46E + 04	8.01E + 06
GER	45	4.92E + 05	8.87E + 03	5.81E + 04	3.90E + 03	1.52E + 04	2.27E + 03	4.83E + 05
ITA	45	3.48E + 06	5.13E + 04	3.94E + 05	3.72E + 04	9.67E + 04	1.53E + 04	3.42E + 06
POR	45	2.68E + 05	2.82E + 03	2.28E + 04	1.34E + 03	4.37E + 03	6.95E + 02	2.64E + 05
KOR	45	4.29E + 04	2.83E + 02	2.78E + 03	6.26E + 02	6.18E + 02	1.01E + 02	4.26E + 04
SPA	45	1.87E + 06	4.02E + 04	3.75E + 05	2.50E + 04	9.98E + 04	2.13E + 04	1.82E + 06
SWI	45	1.40E + 05	2.23E + 03	8.92E + 03	8.43E + 02	3.00E + 03	5.40E + 02	1.38E + 05
THA	45	1.23E + 04	5.96E + 01	1.10E + 03	4.48E + 02	2.22E + 02	6.68E + 01	1.22E + 04
USA	45	1.01E + 07	1.36E + 05	1.40E + 06	1.07E + 05	3.51E + 05	6.85E + 04	9.98E + 06
BR	60	9.38E + 06	1.20E + 05	1.01E + 06	4.65E + 04	2.70E + 05	3.46E + 04	8.71E + 06
GER	60	4.92E + 05	8.87E + 03	5.82E + 04	3.90E + 03	1.52E + 04	2.27E + 03	4.83E + 05
ITA	60	3.48E + 06	5.16E + 04	3.95E + 05	3.72E + 04	9.75E + 04	1.53E + 04	3.43E + 06
POR	60	2.68E + 05	2.84E + 03	2.30E + 04	1.34E + 03	4.43E + 03	6.95E + 02	2.65E + 05
KOR	60	4.29E + 04	2.83E + 02	2.78E + 03	6.26E + 02	6.18E + 02	1.01E + 02	4.26E + 04
SPA	60	1.87E + 06	4.04E + 04	3.76E + 05	2.50E + 04	1.01E + 05	2.13E + 04	1.82E + 06
SWI	60	1.40E + 05	2.23E + 03	8.92E + 03	8.43E + 02	3.01E + 03	5.40E + 02	1.38E + 05
THA	60	1.23E + 04	5.96E + 01	1.10E + 03	4.48E + 02	2.22E + 02	6.68E + 01	1.22E + 04
USA	60	1.01E + 07	1.37E + 05	1.40E + 06	1.07E + 05	3.53E + 05	6.85E + 04	9.99E + 06

optimization we predict death values for 15, 30, 45, and 60 days ahead with errors in the order of 5, 10, 20, and 80%, respectively, for all countries; (d) sudden changes in active curves behavior cannot be predicted by the model without coefficients estimated from outside sources.

Our results show that that our model can fit data from several countries, despite obvious different COVID-19 scenarios among them, such as South Korea and Spain for example. In order to do that, among other things, we estimated the infection rate (β) as an important determinant in the growth of the

infected cases mainly in the early stages of the epidemic and a social distancing coefficient (SD) and a protective coefficient (α) that can cause decreases in rate of transmission. This estimation process provides information to compare different social distance measures adopted among several countries. South Korea results, for instance, exhibits decreased effective transmission rate $\beta(1 - SD)$ compared with other countries and the best social distancing at a rate of 64%. This result concurs with South Korea political decisions (Shin 2020). As our model does not have a quarantined state, the effective testing, contact

Table 5 Percentage errors comparing model future projections of 15, 30, 45, and 60 days with real data for the day of the analysis (May 18, 2020), considering minimizing both active cases and deaths RMSE

Country	Optz	Days	Cases	Deaths	Abs C	Abs D	Days	Cases	Deaths	Abs C	Abs D
BR	J	0	-0.81	-0.03	0.81	0.03	15	-0.82	0.07	0.82	0.07
GER	J	0	-0.37	0.10	0.37	0.10	15	-0.24	0.50	0.24	0.50
ITA	J	0	-0.88	0.37	0.88	0.37	15	-0.74	0.52	0.74	0.52
POR	J	0	-0.39	0.35	0.39	0.35	15	-0.49	0.43	0.49	0.43
KOR	J	0	-0.99	0.01	0.99	0.01	15	-0.99	-0.01	0.99	0.01
SPA	J	0	-0.53	0.61	0.53	0.61	15	-0.48	0.46	0.48	0.46
SWI	J	0	-0.73	0.12	0.73	0.12	15	-0.88	0.03	0.88	0.03
THA	J	0	-0.96	0.02	0.96	0.02	15	-0.96	0.15	0.96	0.15
USA	J	0	-0.95	0.04	0.95	0.04	15	-0.58	0.63	0.58	0.63
		μ	-0.73	0.18	<i>0.73</i>	<i>0.18</i>		-0.69	0.31	<i>0.69</i>	<i>0.31</i>
		STD	0.25	0.22	<i>0.25</i>	<i>0.21</i>		0.26	0.25	<i>0.26</i>	<i>0.24</i>
BR	J	30	-0.90	0.08	0.90	0.08	45	-0.80	0.23	0.80	0.23
GER	J	30	-0.66	0.30	0.66	0.30	45	-0.88	-0.05	0.88	0.05
ITA	J	30	-0.87	1.11	0.87	1.11	45	-0.77	1.21	0.77	1.21
POR	J	30	-0.49	0.86	0.49	0.86	45	-0.40	0.78	0.40	0.78
KOR	J	30	-0.99	0.10	0.99	0.10	45	-0.99	0.05	0.99	0.05
SPA	J	30	-0.33	1.10	0.33	1.10	45	-0.74	1.36	0.74	1.36
SWI	J	30	-0.87	0.20	0.87	0.20	45	-0.79	0.42	0.79	0.42
THA	J	30	-0.97	0.21	0.97	0.21	45	-0.96	0.28	0.96	0.28
USA	J	30	-0.95	1.54	0.95	1.54	45	-0.87	1.15	0.87	1.15
		μ	-0.78	0.61	<i>0.78</i>	<i>0.61</i>		-0.80	0.60	<i>0.80</i>	<i>0.61</i>
		STD	0.23	0.55	<i>0.23</i>	<i>0.55</i>		0.17	0.53	<i>0.17</i>	<i>0.52</i>
BR	J	60	-0.69	0.40	0.69	0.40	30*	0.44	0.86	0.44	0.86
GER	J	60	-0.99	-0.67	0.99	0.67	30*	-0.26	0.39	0.26	0.39
ITA	J	60	-0.87	2.17	0.87	2.17	30*	-1.32	-0.06	1.32	0.06
POR	J	60	-0.82	1.67	0.82	1.67	30*	-1.76	-0.40	1.76	0.40
KOR	J	60	-1.00	0.04	1.00	0.04	30*	-0.08	0.10	0.08	0.10
SPA	J	60	-0.77	2.50	0.77	2.50	30*	-1.30	0.29	1.30	0.29
SWI	J	60	-0.05	1.03	0.05	1.03	30*	-0.73	0.31	0.73	0.31
THA	J	60	-0.74	1.18	0.74	1.18	30*	-0.89	-0.38	0.89	0.38
USA	J	60	-0.91	2.12	0.91	2.12	30*	-1.43	0.20	1.43	0.20
		μ	-0.76	1.16	<i>0.76</i>	<i>1.31</i>		-0.81	0.15	<i>0.91</i>	<i>0.33</i>
		STD	0.29	1.07	<i>0.29</i>	<i>0.86</i>		0.73	0.39	<i>0.58</i>	<i>0.24</i>

*30-day percentage errors in futures projections considering data from March 19, 2020 to predict April 18, 2020. μ represents the mean across countries and STD the standard deviation. Italics emphasize results related to the choice of optimization.

tracing, and quarantining implemented by Korea was reflected not only in a greater SD values but also an increased protection rate of $\alpha = 0.036$. The worse protection rates were found for Brazil ($\alpha = 0.016$) and the USA ($\alpha = 0.02$) most likely caused by poor political decision and downplaying by officials of the seriousness of the virus in the beginning of the crisis (Abutaleb et al. 2020; Andreoni 2020).

Furthermore, in order to adequately model, countries where the number of deaths are critically above the expected number considering COVID-19 death mortality rates even considering possible age effects (Li et al. 2020; WHO 2020), we introduced a coefficient l to the model.

This coefficient represents the percentage of people that went from infectious to death without access to hospital care. Introducing l was a novel idea in SEIR model studies. It was done to account for the sad reality that many people are facing during the COVID-19 pandemic, as many people have passed away for the lack of available ICU and/or hospital beds, especially in some regions where the outbreak was not early contained, Italy for example (Tondo 2020). Nevertheless, in order to accurately estimate the value of l , one need to know c (specific parameter to account the fraction of hospitalized that becomes critical cases) and f

Table 6 Percentage errors comparing model future projections of 15, 30, 45, and 60 days to real data for the day of the analysis (May 18, 2020), considering minimizing deaths RMSE

Country	Optz	Days	Cases	Deaths	Abs C	Abs D	Days	Cases	Deaths	Abs C	Abs D
BR	D	0	-0.84	0.02	0.84	0.02	15	-0.86	0.05	0.86	0.05
GER	D	0	-0.73	0.00	0.73	0.00	15	-0.76	0.05	0.76	0.05
ITA	D	0	-0.99	0.07	0.99	0.07	15	-0.99	0.10	0.99	0.10
POR	D	0	-0.92	0.08	0.92	0.08	15	-0.99	0.04	0.99	0.04
KOR	D	0	-1.00	-0.03	1.00	0.03	15	-1.00	-0.03	1.00	0.03
SPA	D	0	-0.99	-0.02	0.99	0.02	15	-0.99	-0.05	0.99	0.05
SWI	D	0	-0.85	0.04	0.85	0.04	15	-0.92	0.04	0.92	0.04
THA	D	0	-0.97	0.04	0.97	0.04	15	-0.99	0.04	0.99	0.04
USA	D	0	-0.98	-0.01	0.98	0.01	15	-0.93	0.01	0.93	0.01
		μ	-0.92	0.02	0.92	<i>0.03</i>		-0.94	0.03	0.94	<i>0.05</i>
		STD	0.09	0.04	0.09	<i>0.03</i>		0.08	0.04	0.08	<i>0.02</i>
BR	D	30	-0.91	0.07	0.91	0.07	45	-0.65	0.40	0.65	0.40
GER	D	30	-0.80	0.08	0.80	0.08	45	-0.88	-0.16	0.88	0.16
ITA	D	30	-1.00	0.09	1.00	0.09	45	-0.96	0.43	0.96	0.43
POR	D	30	-0.99	0.12	0.99	0.12	45	-0.97	0.20	0.97	0.20
KOR	D	30	-1.00	-0.04	1.00	0.04	45	-1.00	-0.02	1.00	0.02
SPA	D	30	-1.00	-0.07	1.00	0.07	45	-0.99	0.13	0.99	0.13
SWI	D	30	-0.97	0.09	0.97	0.09	45	-0.91	0.15	0.91	0.15
THA	D	30	-0.99	0.10	0.99	0.10	45	-0.95	0.30	0.95	0.30
USA	D	30	-0.99	-0.02	0.99	0.02	45	-0.97	0.18	0.97	0.18
		μ	-0.96	0.05	0.96	<i>0.08</i>		-0.92	0.18	0.92	<i>0.22</i>
		STD	0.07	0.07	0.07	<i>0.03</i>		0.11	0.19	0.11	<i>0.13</i>
BR	D	60	0.63	1.46	0.63	1.46	30*	0.10	0.69	0.10	0.69
GER	D	60	-0.97	-0.36	0.97	0.36	30*	-0.92	-0.33	0.92	0.33
ITA	D	60	-0.98	0.73	0.98	0.73	30*	-0.32	0.46	0.32	0.46
POR	D	60	-0.97	0.36	0.97	0.36	30*	-0.59	0.43	0.59	0.43
KOR	D	60	-1.00	0.01	1.00	0.01	30*	-0.06	0.15	0.06	0.15
SPA	D	60	-0.98	0.62	0.98	0.62	30*	-0.48	0.31	0.48	0.31
SWI	D	60	-0.64	0.61	0.64	0.61	30*	-0.48	0.13	0.48	0.13
THA	D	60	-0.69	1.37	0.69	1.37	30*	-1.03	-0.59	1.03	0.59
USA	D	60	-0.94	1.78	0.94	1.78	30*	-1.28	0.11	1.28	0.11
		μ	-0.73	0.73	0.87	<i>0.81</i>		-0.56	0.15	0.58	<i>0.35</i>
		STD	0.52	0.70	0.16	<i>0.59</i>		0.45	0.40	0.42	<i>0.21</i>

*30-day percentage errors in futures projections considering data from March 19, 2020 to predict April 18, 2020. μ represents the mean across countries and STD the standard deviation. Italics emphasize results related to the choice of optimization.

(specific parameter to account the fraction of people in critical state that died) from outside sources.

Our model predicted a basic reproduction number R_0 of 2.24 (0.52). The basic reproduction number represents the average number of secondary cases that result from the introduction of a single infectious case in a susceptible population (Anastassopoulou et al. 2020). Considering the importance of such parameter, several other papers have tried using different methods to estimate this parameter for COVID-19, and our values fall within the range of values reported so far. In their review, Liu et al. (2020) reported two studies using stochastic methods that estimated R_0 ranging from 2.2 to 2.68, six

studies, where mathematical methods, with results ranging of 1.5 to 6.49, and finally three studies that used statistical methods such as exponential growth with estimations ranging from 2.2 to 3.58.

Additionally, we found a worldwide mean of latent period of 1 (0.2) and infectious period of approximately 8.9 (1.5) days. The mean estimated latent period found here is smaller than some previously reported, such as in PENG et al. (2020) and GUAN et al. (2020) who reported estimates the latent median times around 2–3 days. Nevertheless, our results corroborate with the idea that COVID-19 transmission may occur in the pre-symptomatic phase and that COVID-19 patients

Table 7 Percentage errors comparing model future projections of 15, 30, 45, and 60 days to real data for the day of the analysis (May 18, 2020), considering minimizing active cases RMSE

Country	Optz	Days	Cases	Deaths	Abs C	Abs D	Days	Cases	Deaths	Abs C	Abs D
BR	C	0	-0.80	-0.05	0.80	0.05	15	-0.82	0.14	0.82	0.14
GER	C	0	-0.37	0.10	0.37	0.10	15	-0.24	0.98	0.24	0.98
ITA	C	0	-0.88	0.37	0.88	0.37	15	-0.80	0.81	0.80	0.81
POR	C	0	-0.39	0.35	0.39	0.35	15	-0.49	0.43	0.49	0.43
KOR	C	0	-0.98	1.21	0.98	1.21	15	-0.99	0.07	0.99	0.07
SPA	C	0	-0.52	0.90	0.52	0.90	15	-0.48	0.46	0.48	0.46
SWI	C	0	-0.73	0.12	0.73	0.12	15	-0.88	0.03	0.88	0.03
THA	C	0	-0.96	0.15	0.96	0.15	15	-0.96	0.15	0.96	0.15
USA	C	0	-0.94	0.77	0.94	0.77	15	-0.58	0.63	0.58	0.63
		μ	-0.73	0.44	0.73	0.45		-0.69	0.41	0.69	0.41
		STD	0.24	0.43	0.24	0.41		0.26	0.34	0.26	0.34
BR	C	30	-0.88	0.46	0.88	0.46	45	-0.80	0.17	0.80	0.17
GER	C	30	-0.66	0.30	0.66	0.30	45	-0.88	-0.09	0.88	0.09
ITA	C	30	-0.86	1.66	0.86	1.66	45	-0.75	1.24	0.75	1.24
POR	C	30	-0.48	1.28	0.48	1.28	45	-0.40	0.76	0.40	0.76
KOR	C	30	-0.98	2.04	0.98	2.04	45	-0.98	0.04	0.98	0.04
SPA	C	30	-0.33	1.50	0.33	1.50	45	-0.73	1.46	0.73	1.46
SWI	C	30	-0.87	0.20	0.87	0.20	45	-0.78	0.54	0.78	0.54
THA	C	30	-0.99	1.40	0.99	1.40	45	-0.96	0.28	0.96	0.28
USA	C	30	-0.93	3.07	0.93	3.07	45	-0.87	1.19	0.87	1.19
		μ	-0.78	1.32	0.78	1.32		-0.80	0.62	0.80	0.64
		STD	0.23	0.92	0.23	0.92		0.17	0.57	0.17	0.55
BR	C	60	-0.69	0.44	0.69	0.44	30*	0.86	0.44	0.86	0.44
GER	C	60	-0.99	-0.67	0.99	0.67	30*	0.39	-0.30	0.39	0.30
ITA	C	60	-0.87	2.17	0.87	2.17	30*	-0.06	-1.32	0.06	1.32
POR	C	60	-0.82	1.67	0.82	1.67	30*	-0.40	-1.76	0.40	1.76
KOR	C	60	-0.99	1.53	0.99	1.53	30*	0.20	-1.60	0.20	1.60
SPA	C	60	-0.77	2.50	0.77	2.50	30*	0.29	-1.30	0.29	1.30
SWI	C	60	-0.06	1.09	0.06	1.09	30*	0.31	-0.78	0.31	0.78
THA	C	60	-0.83	1.12	0.83	1.12	30*	-0.41	-0.92	0.41	0.92
USA	C	60	-0.91	1.69	0.91	1.69	30*	0.20	-1.09	0.20	1.09
		μ	-0.77	1.28	0.77	1.43		0.16	-0.96	0.35	1.06
		STD	0.28	0.95	0.28	0.67		0.40	0.69	0.23	0.50

* 30-day percentage errors in futures projections considering data from March 19, 2020 to predict April 18, 2020. μ represents the mean across countries and STD the standard deviation. Italics emphasize results related to the choice of optimization.

may have an inconsiderable latent non-infectious period. The mean infectious period of 9 days is within expected range estimated by recent publications (Guo et al. 2020; Hou et al. 2020).

Our results indicate that, despite all uncertainty and biases in the data collected, lack of testing in several countries and possible changes in policies and people's behavior regarding the COVID-19 our proposed mathematical modeling may help predict 15 days ahead values of total deaths with errors in the order of 5% and 30 days ahead values of active cases with errors in the order of 30%. Moreover, a reliable 2-week prediction of the number of deaths suggests that the model

may also be used to determine the number of hospital and ICU beds that a region will need ahead of time enough for people to prepare themselves for it. Unfortunately, we could not get reliable data of number of hospitalizations and ICU patients in the different countries studied here to verify the certainty of our predictions for the values estimated by the model, and we urge future research to do so. Furthermore, future application of our model should consider including stratification by age groups (Li et al. 2020) and coefficients to account for temperature variations and people's density (Chen et al. 2020; Wang et al. 2020).

Additionally, for Brazil where the active cases are still fast-growing errors in prediction can be large (Fig. 4). The larger errors in such cases happened because there is less data for the optimization process to fit the data to the models' parameters and the fact the active cases and accumulated death curves are still, approximately, exponentially growing (Ranjan 2020). Because of the simplicity of the curve, different optimization solutions can fit the data but yield quite different future projections. For example, different combinations of α and β may cause similar behavior patterns for the beginning of the curve. Our results are in agreement with recent study by Ranjan (2020), who adds that modeling of an epidemic during its progress is very challenging as the parameters such as transmission rate and basic reproduction number are different for different geographical regions and depend on many social and environmental factors. They also concluded that the early stage of an epidemic is relatively easy to model and the modeling of later stages to predict the decline and eventual flattening of the curve is very challenging as more known parameters need to be included in the model. The inclusion of effects due to isolation and quarantine adds to the complication. Although technically we solved this issue by including in our model three time-changing coefficients α , SD , and l , they are hard to find by optimization for countries in the beginning stages. This happens mainly because SD and l are time-dependent triggered and the optimization process attributes random values for both these coefficients and their time "activations." With larger t_{sd} and t_l than current time, different values of SD and l can yield the same temporal trends for the beginning of the curves but significantly different behaviors after times t_{sd} and t_l . In other words, in countries where the epidemic is still in its pre-peak stages, especially during the fast-initial growing phase, some of the model coefficients, especially SD , α , t_{sd} , and t_l , should be estimated from outside sources and/or used to infer possible future scenarios dependent upon future defined policies, such as, for example, an enforcement of social distance measures. Furthermore, sudden changes in SD after a period may also cause a rapid increase in active cases that cannot be predicted by the model (e.g., Figure 4; Spain, and Portugal); to predict such cases, data from SD and ideally protection rates should be obtained from outside sources. Studies to test such hypothesis should be made.

Conclusion

In response to the rapid global dissemination of the COVID-19, on the 11th of March the WHO has declared the outbreak a pandemic motivating further research in epidemiological mathematical modelling. The results suggest that the presented model may be used to predict 15 days ahead values of total deaths with errors in the order of 5%. These errors may be minimized if social distance data are inputted into the model.

Sudden changes in social distance measures could not be predicted by the model using optimization alone.

Acknowledgments The authors would like to thank Dr. Osmar Pinto JR, Dra. Iara RCA Pinto, Emely Flores, Leandro Dalmarco, Fernando Torres Balbina, Fabricio Duarte, Henrique Touguinha, Guilherme Ferro, Marco Antonio Ridenti, and all others who have helped us by sharing data and information during this project.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- Abutaleb Y, et al. The U.S. was beset by denial and dysfunction as the coronavirus raged. The Washington Post. Washington, Apr., 4, 2020. Available in: <https://www.washingtonpost.com/national-security/2020/04/04/coronavirus-government-dysfunction/?arc404=true>, accessed in Apr.16, 2020. 9:55.
- Anastassopoulou C, et al. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. PLoS One. 2020;15(3):e0230405.
- Andreoni, M. Coronavirus in Brazil: What You Need to Know. New York Times. New York, July, 2, 2020. Available in: <https://www.nytimes.com/article/brazil-coronavirus-cases.html> accessed in: Jul. 7, 2020. 07:07.
- Chen, B. et al. Roles of meteorological conditions in COVID-19 transmission on a worldwide scale. medRxiv, v. 11, p. 2020.03.16.20037168, 2020.
- Ferguson, N. M. et al. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College, March, 2020.
- Guan W, et al. Clinical characteristics of 2019 novel coronavirus infection in China. medrxiv. 2020;34:2.
- Gudbjartsson DF, Helgason A, Jonsson AH, et al. Early spread of SARS-Cov-2 in the Icelandic population. N Engl J Med. 2020;382:1–34. <https://doi.org/10.1056/NEJMoa2006100>.
- Guo YR, et al. The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak – an update on the status. Military Medical Research. 2020;7(1):1–10.
- Hasell, J. et al. To understand the global pandemic, we need global testing—the Our World in Data COVID-19 Testing dataset. Our World in Data, Mar 31, 2020. Coronavirus Disease (COVID-19). Available in: <https://ourworldindata.org/covid-testing>, accessed in: Apr. 17, 2020. 19:07.
- Hou, C. et al. The effectiveness of the quarantine of Wuhan city against the Corona Virus Disease 2019 (COVID-19): well-mixed SEIR model analysis. J Med Virol, p. 0–3, 2020.
- Kennedy DM, Zambrano GJ, Wang Y, Neto OP. Modeling the effects of intervention strategies on COVID-19 transmission dynamics. J Clin Virol. 2020;128:104440. <https://doi.org/10.1016/j.jcv.2020.104440>.
- Li, X. et al. Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. J Allergy Clin Immunol, 2020.
- Liu Y, et al. The reproductive number of COVID-19 is higher compared to SARS coronavirus. Journal of travel medicine. 2020;27(2):1–4.
- Peng, L. et al. Epidemic analysis of COVID-19 in China by dynamical modeling. In press, n. 2020.
- Ranjan R. Estimating the final epidemic size for COVID-19. medRxiv. 2020;1:1–15.

- Shin, H. South Korea extends intensive social distancing to reach 50 daily coronavirus cases. REUTERS. New York, April 3, 2020. World News. Available: <https://www.reuters.com/article/us-health-coronavirus-southkorea/south-korea-extends-intensive-social-distancing-to-reach-50-daily-coronavirus-cases-idUSKBN21M02P>, accessed in April 17, 2020, 09:45.
- Tondo, L. Italian hospitals short of beds as coronavirus death toll jumps. The Guardian. Palermo, Mar.09, 2020. Available in: <https://www.theguardian.com/world/2020/mar/09/italian-hospitals-short-beds-coronavirus-death-toll-jumps>, accessed in: Apr., 17, 2020. 11:00.
- Verity R, Dorigatti I, Winskill P, et al. Estimates of the severity of coronavirus deases 2019: a model-based analysis. Lancet Infect Dis. 2020. [https://doi.org/10.1016/S1473-3099\(20\)30243-7](https://doi.org/10.1016/S1473-3099(20)30243-7).
- Walker, PGT. et al. The global impact of COVID-19 and strategies for mitigation and suppression. **Imperial College**, n. March, p. 1–19, 2020.
- Wang, J. et al. High temperature and high humidity reduce the transmission of COVID-19. SSRN Electron J, 2020.
- WHO Director-General's opening remarks at the media briefing on COVID-19. World Health Organization (WHO) 2020. Available in: <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19%2D%2D-30-march-2020>, accessed in Apr. 17, 2020. 10:40.
- Wu, J. T. et al. Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. Nature Medicine, p. 1–5, 2020.

Consulted links

- APMonitor Optimization Suite. COVID-19 optimal control response. 2020. Available: <https://apmonitor.com/do/index.php/Main/COVID-19Response>, accessed in Apr. 18, 2020. 17:29.
- UFPEL Epidemiologia da COVID-19 no Rio Grande do Sul, EPICOVID-19. Pelotas, Apr. 15, 2020. Available in: <https://ccs2.ufpel.edu.br/wp/2020/04/15/ufpel-apresenta-primeiros-resultados-do-estudo-sobre-covid-19-no-rs/>, accessed in: Apr. 17, 2020. 19:22.
- University of Basel. COVID-19scenarios developed at the University of Basel. ©2020. Available in: <https://covid19-scenarios.org/about> , accessed in Apr. 18, 2020. 17:247.
- WorldoMeter. COVID-19, 2020. Available: <https://www.worldometers.info/coronavirus/>, accessed in 13/04, 21:45.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.