Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

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

Osmar Pinto Neto (osmarpintoneto@hotmail.com)
Anhembi Morumbi https://orcid.org/0000-0002-3978-8320

José Clark Reis
Arena235 Research Lab

Ana Carolina Brisola Brizzi
Anhembi Morumbi University

Gustavo José Zambrano
Arena235 Research Lab

Joabe Marcos de Souza
Universidade de São Paulo

Wellington Pedroso Estevam de Amorim
Anhembi Morumbi University

Rodrigo Cunha de Mello Pedreiro
Anhembi Morumbi University

Bruno de Matos Brizzi
Arena235 Research Lab

Ellysson Oliveira Abinader
Instituo Abinader

Renato Amaro Zangaro
Anhembi Morumbi

Research Article

Keywords: COVID-19, SEIR model, pandemic, mathematical modelling, countries, coronavirus, optimization.

Posted Date: May 1st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-26264/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

April 18, 2020

Submitted to Research on Biomedical Engineering Journal

"Special Issue: Emerging technologies for fighting COVID-19"

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

Osmar Pinto Neto[1,2,4*], José Clark Reis[2], Ana Carolina Brisola Brizzi[1,2], Gustavo José Zambrano[2], Joabe Marcos de Souza[2,3], Wellington Amorim Pedroso[1,2], Rodrigo Cunha de Mello Pedreiro[1,6,7], Bruno de Matos Brizzi[2], Ellysson Oliveira Abinader[5], Renato Amaro Zângaro[1,4].

[1] Anhembi Morumbi University, Biomedical Engineering Department, São Paulo, SP, Brazil.

[2] Arena235 Research Lab – São José dos Campos, SP, Brazil.

[3] Universidade de São Paulo, Departamento de Engenharia Aeronáutica, São Paulo, SP, Brazil.

[4] Center for Innovation, Technology and Education – CITE, Parque Tecnológico de São José dos Campos, São José dos Campos, SP, Brazil.

[5] Instituto Abinader, Manaus , AM, Brazil.

[6] Estácio de Sá University, Nova Fribugo, RJ, Brazil.

[7] Santo Antônio de Pádua College, Santo Antônio de Pádua, RJ, Brasil,

*

Corresponding Author:

Osmar Pinto Neto

osmarpintoneto@hotmail.com

Parque Tecnológico de São José dos Campos, Estrada Dr. Altino Bondensan, 500, São José dos Campos, SP, Brazil, 12247-016.

Abstract

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. In light of 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 a given social, economic and political scenario. We present a novel generalized Susceptible-Exposed-Infected-Recovered (SEIR) compartmental model and test it using a global optimization algorithm with data collected from the WHO. Our main results were: (a) our model was able to accurately fit the data of all countries tested (b) it is possible to predict values for one week ahead with errors in the order of 15% for the number of cases and 30% in the number of deaths for all countries; (c) predictions are better for countries where the active cases curve already reached the maximum; the error being in the order of 10% in the number of cases and 20% in the number of deaths; (d) for countries where the active curve is still growing, different optimization solutions can be found that fit the data; so, to predict future behavior in this scenarios some of the model coefficients should be estimated from outside sources or based on generalized results from other countries according to their health policies of social distance, quarantining and case test and tracing.

Keywords: COVID-19; SEIR model; pandemic; mathematical modelling; countries; coronavirus; optimization.

Introduction

In December 2019, in China, a series of atypical pneumonia cases have emerged caused by a new Coronavirus, previously called 2019-nCoV or SARS-CoV-2 by the International Committee on Taxonomy of Viruses (ICTV), 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 3,341 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 epidemiological investigations, the period of incubation is 1 to 14 days, more evidently between 3 to 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 to other Coronavirus infections and, to better understand it, it's 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 Unities (UCI). 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 has 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 the SEIR model in order to better represent the behavior of COVID-19 and better predict the outcome of the pandemic in a given social, economic and political canvas. Moreover, a higher fidelity 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, ventilators and ICU's will be needed at a certain point in time to maximize the usage of public resources. A higher fidelity model will also lay the foundation and spark a necessary discussion to prepare authorities to new outbreaks in the future.

Methodology

The SUEIHCDR model

We present a generalized SEIR compartmental model using novel and recently suggested ideas and concepts (APMonitor Optimization Suite; PENG et al., 2020, University of Basel). It is composed of eight compartments: Susceptible, Unsusceptible, Exposed, Infected, Hospitalized, Critical, Dead, and Recovered (SUEIHCDR, Figure 1).

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

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

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

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 (Equation 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 (Equation 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 (Equation 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.

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

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/γ (Equation 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/δ (Equation 6) days and can have three different outcomes. Considering *m* as a specific parameter to account the fraction of infectious that are 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 (Equation 7) so that the time when hospital bed became unavailable could be modeled (t_m), as well as the duration that hospital were full (*dur*).

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

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

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

Hospitalized people (Equation 8) stay hospitalized for 1/ζ 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/ ε in the ICU (Equation 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 (Equation 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*).

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

At last, death (Equation 11) arises either from lack of available treatment (*l*), or from critical cases in the ICU (*f*).

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

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.R14a). 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, Argentina, Chile, Spain, Italy, South Korea, Portugal, Switzerland, Thailand and United States. Lack of testing and under-notification of active cases has been largely reported for the covid-

19 (HASELL et al., 2020; Wordometers.info; UFPel); in consequence, active cases data were corrected considering the number of tests taken in each country per habitants in proportion to the number of tests taken in Switzerland, at the present moment, the country with the greatest amount of tests per million people taken, and assuming a 20% increase in cases as the proportion of tests grows. Argentina received the greatest multiplying factor of 11.

We used a custom build MATLAB global optimization algorithm using Monte Carlo iterations 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), minimizing a goal function (*J*) as a combination of Active Cases and Death time series (Equation 12).

$$\begin{cases} J = (1-p) * RMSE(Active\ Cases) + p * RMSE(Deaths) \\ p = \frac{mean(Active\ Cases)/mean(Death)}{\frac{mean(Active\ Cases)}{mean(Death)}+1} \end{cases} \quad (12)$$

Corrected active cases curves were fitted to the sum of the models infected, hospitalized and critical cases time series. Data under 50 active cases were discarded. Initial values for each compartmental parameter had ranges proportional to the following initial values (Table 1): infected initial values (*I0*) were determined as the corrected actives cases first value greater than 50; exposed initial values (*E0*) were 0.5 x I0; hospitalized initial values (*H0*) were 0.2 x I0; critical cases initial values (*C0*) were 0.7 x I0; deaths initial values (*D0*) were obtained from the accumulated deaths real data; similarly, recovered initial values (*Re0)* were obtained from the recovered real data. Results are presented as mean (standard deviation). The model results presented are based on an average of 1250 runs.

Results

Our model was able to accurately fit the data of all countries tested independent of what stage of the epidemic they were. Tables 2 and 3 present the optimized inputs for April 14, 2020 and April 07, 2020 for Germany, Argentina, Chile, Spain, Italy, South Korea, Portugal, Switzerland, Thailand and United States, respectively for each country. Note that Table 3 shows two possible optimizations results for Portugal. For the most recent date we have that: mean protective rate (α) was 0.05 (0.02); mean infectious rate (β) was 0.68 (0.07); mean fraction of infectious that are asymptomatic or mild (m) was 0.76 (0.03); the mean fraction of infectious people that died with no treatment (l) was 0.01 (0.01); the fraction of severe cases that turn critical (c) was 0.27 (0.09); the mean fraction of critical cases that are fatal (f) was 0.45 (0.04) and the mean social distancing parameter (SD) was 0.36 (0.12). Table 4 shows the inverse values of γ, δ, ζ, ε: the mean latent period was 0.4 (0.1) days; the mean infectious period was 8.7 (2.1); the mean hospitalized period was 3.7 (0.2) days and the mean period in ICU was 10.9 (2.6) days. The reproductive number (**R0**) estimated from the model is also presented in Table 4, mean **R0** found was 3 (0.8).

Figures 2, 3, 4 and 5 show the model results for Spain, South Korea, United States and Portugal. Note that Figure 5 shows two possible optimizations for Portugal using data from April 07, 2020. It shows two different possible solutions for the optimization problem leading to two different projected outcomes.

Table 5 shows the model results for total number of infected, hospitalized, ICU patients, and deaths, as well as peak day hospitalization, peak day ICU, number of deaths for lack of treatment and percentage of death caused by lack of treatment. Results indicate deaths in the thousands for every country but Chile, Korea and Thailand. United States has a peak day of more than 80 thousand hospitalized patients. Additionally, Spain projects more than 1 million recovered people. Italy has, by far, the greatest projected percentage of deaths coming from

lack of treatment, with 35%, followed by Argentina with 30%. According to the model estimations, Argentina will have 6176 ICU patients treated by the end of the epidemic. The model suggests that South Korea and Thailand will have no deaths for lack hospitalization and peaks in ICU patients in one day fewer than 200.

Finally, Table 6 shows the percentage errors in estimation for 7 days ahead considering the model run for April 7 and comparing to real data for April 14. The results indicate a mean error in cases estimation of 0.15 (0.08) and 0.28 (0.14) for deaths. Errors were the greatest for the United States. Errors for countries that either have already peaked or a very close to the peak in active cases (Spain, Italy, South Korea, Thailand) were in general smaller: mean of 0.09 (0.21) for cases and 0.21 (0.12) for deaths.

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. We tested our model using a global optimization algorithm and data collected from the WHO. Our main findings were: (a) our model was able to accurately fit the data of all countries tested independent of what stage of the epidemic they were using optimized coefficient values in agreement with recent reports; (b) using our model, large ranges for each input, and optimization we predict values for one week ahead with errors in the order of 15% for the number of cases and 30% in the number of deaths for all countries; (c) predictions are better for countries where the active cases curve already reached the maximum, in the order of 10% in the number of cases and 20% in the number of deaths; (d) for countries where the active curve is still growing, different optimization solutions can be found tot fit the data; so, to predict future behavior in this scenarios, some of the model coefficients should be estimated from

outside sources or based on generalized results from other countries according to their health policies of social distance, quarantining and case test and tracing.

Our results show that that our model can fit data from several countries, despite obvious very 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 β (1-SD) compared to other countries and the best social distancing at a rate of 60%. This result concurs with South Korea political decisions (SHIN et al., 2020). As our model does not have a quarantined state, the effective testing, contact tracing and quarantining implemented by Korea was reflected not only in a greater SD values but also an increased protection rate of α=0.08, only matched by Chile. The worse protection rates were found for Italy (α =0.03) and the United States (α =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 (HOROWITZ et al., 2020, ABUTALEB et al., 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, World Health Organization, 2020), we introduced a coefficient l to the model. This coefficient represent the percentage of people that went from infectious to death without access to hospital care. Introducing l was a novel idea in SEIR models 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).

Our model predicted a basic reproduction number **R0** of 3 (0.8). 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 all., 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 **R0** 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 0.4 (0.1) and infectious period of approximately 9 (2) 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 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 within a week's period, our proposed mathematical modelling may help predict values of corrected active cases and totals deaths with errors in the order of 15% for the number of active cases and 30% in the number of deaths. Moreover, a reliable one-week prediction of the number of active cases and accumulated 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, we found that predictions are better for countries where the active cases curve already reached the maximum; in such cases, the one-week estimation error was approximately 10% in the number of cases and 20% in the number of deaths. For countries where the active cases is still fast-growing errors were larger, and in the case of the United States were found to be 30% for the cases and 50% for the deaths. 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 deaths curves are still, approximately, exponentially growing (RANJAN, 2020). Because of the simplicity of the curve, different optimization solutions can fit the data but yield very different future projections (see Figure 5). 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, basic reproduction number etc. 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 were 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.

Conclusion

In response to the rapid global dissemination of the virus, on the 11th of Mars the WHO has declared the outbreak a pandemic motivating research to improve the current SEIR models in order to better represent and predict the behavior of COVID-19. Our main results suggest that mathematical modelling may help predict on average one-week ahead values of corrected active cases with errors of 30% and totals deaths in the order of 15%. For countries after the peak stage of the epidemic, the prediction errors were 10% in the number of cases and 20% in the number of deaths.

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.

Conflict of interest statement

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Funding Statement

The authors received no external funding for this research.

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**, v. 15, n. 3, p. e0230405, 2020.

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**, v. 34, n. 2, 2020.

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**, v. 7, n. 1, p. 1–10, 2020.

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. **Journal of Medical Virology**, p. 0–3, 2020.

HOROWITZ, J.; BUBOLA, E.; POVOLEDO, E. Italy, Pandemic's New Epicenter, Has Lessons for the World. **New York Time**. New York, Mar., 21, 2020. Europe. Available in: https://www.nytimes.com/2020/03/21/world/europe/italy-coronavirus-center-lessons.html, accessed in Apr., 17, 2020. 10:05.

LI, X. et al. Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. The Journal of allergy and clinical immunology, 2020.

LIU, Y. et al. The reproductive number of COVID-19 is higher compared to SARS coronavirus. **Journal of travel medicine**, v. 27, n. 2, p. 1–4, 2020.

PENG, L. et al. Epidemic analysis of COVID-19 in China by dynamical modeling. ***In press***, n. February, 2020.

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.

RANJAN, R. Estimating the final epidemic size for COVID-19. **medRxiv**, v. 1, p. 1–15, 2020.

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.

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 Electronic Journal**, 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---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.

Centro of Disease Control and Prevention (CDC). Non-Pharmaceutical Interventions (NPIs), 2019. Available: http//www.cdc.gov/nonpharmaceuticalinterventions/community/index.html, accessed in 13/04, 23:20.

Our World in Data. COVID-19, 2020. Available: https://ourworldindata.org/search?q=covid+19, accessed in 13/04, 23:53.

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.

Table Legends:

Table 1: Input coefficients to the model and respective ranges.

Table 2: Optimized inputs for Germany, Argentina, Chile, Spain, Italy, South Korea, Portugal, Switzerland, Thailand and United States, respectively, for April 14, 2020. μ stands for the mean across countries, and STD for the standard deviation.

Table 3: Optimized inputs for Germany, Argentina, Chile, Spain, Italy, South Korea, Portugal, Switzerland, Thailand and United States, respectively, for April 07, 2020. μ stands for the mean across countries, and STD for the standard deviation.

Table 4: Inverse of the model optimized coefficients of γ, δ, ζ, ε representing latent, infectious, hospitalization and critical cases mean duration in days, as well as the model estimated basic reproductive number (**R0**) for April 14, 2020 for Germany, Argentina, Chile, Spain, Italy, South Korea, Portugal, Switzerland, Thailand and United States, respectively. μ stands for the mean across countries, and STD for the standard deviation.

Table 5: Model results for total number of infected, hospitalized, ICU patients, and deaths, as well as peak day hospitalization, peak day ICU, number of deaths with no treatment and percentage of death caused by lack of treatment for Germany, Argentina, Chile, Spain, Italy, South Korea, Portugal, Switzerland, Thailand and United States, respectively. μ stands for the mean across countries, and STD for the standard deviation.

Table 6: Seven days percentage errors and absolute percentage errors in estimations for the active cases (C) and cumulated deaths (D), considering the model run for April 7 and comparing to real data for April 14 for Germany, Argentina, Chile, Spain, Italy, South Korea, Portugal, Switzerland, Thailand and United States, respectively. μ stands for the mean across countries, and STD for the standard deviation.

Figure Legends

Figure 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.

Figure 2: Model results for active cases (a) and accumulated deaths (b) for Spain, using data from April 07, 2020 and April 14, 2020.

Figure 3: Model results for active cases (a) and accumulated deaths (b) for South Korea, using data from April 07, 2020 and April 14, 2020.

Figure 4: Model results for active cases (a) and accumulated deaths (b) for United States, using data from April 07, 2020 and April 14, 2020.

Figure 5: Model results for active cases (a) and accumulated deaths (b) for Portugal, using data from April 07, 2020 and April 14, 2020. Note that there are two possible optimizations for Portugal using data from April 07, 2020.

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
$Re0$	Re0/2	2Re0
$D0$	D0/2	2E0
SD	0.00	0.75
m_{sd}	0.00	40

Table 2: Optimized inputs for Germany, Argentina, Chile, Spain, Italy, South Korea, Portugal, Switzerland, Thailand and United States, respectively, for April 14, 2020. μ stands for the mean across countries, and STD for the standard deviation.

Inp.	GER	ARG	CHI	SPA	ITA	KOR	POR	SWI	THA	USA	μ	STD
α	0.05	0.03	0.08	0.04	0.03	0.08	0.04	0.05	0.06	0.02	0.05	0.02
β	0.69	0.62	0.63	0.74	0.80	0.75	0.61	0.65	0.61	0.67	0.68	0.07
γ	1.92	3.32	1.67	2.93	3.14	2.12	2.84	2.66	2.02	2.35	2.50	0.56
δ	0.11	0.09	0.09	0.14	0.17	0.09	0.13	0.15	0.11	0.13	0.12	0.03
ζ	0.26	0.27	0.26	0.29	0.28	0.26	0.27	0.25	0.26	0.28	0.27	0.01
ε	0.10	0.08	0.07	0.10	0.12	0.10	0.11	0.10	0.06	0.11	0.10	0.02
m	0.74	0.80	0.79	0.74	0.72	0.81	0.74	0.73	0.76	0.73	0.76	0.03
l	0.011	0.006	0.008	0.015	0.039	0.000	0.016	0.015	0.008	0.014	0.013	0.010
rl	0.09	0.08	0.14	0.11	0.10	0.09	0.09	0.09	0.11	0.09	0.10	0.02
t_l	14	31	30	21	3	27	14	17	23	27	21	9
dur	24	10	26	17	28	26	27	24	23	26	23	5
c	0.23	0.14	0.19	0.34	0.38	0.15	0.33	0.28	0.22	0.40	0.27	0.09
f	0.47	0.40	0.42	0.45	0.55	0.43	0.45	0.48	0.41	0.46	0.45	0.04
$E0$	19	22	21	21	22	20	22	20	21	26	21.40	1.90
$I0$	39	47	50	47	45	41	46	41	42	38	43.60	3.95
$H0$	7	38	40	35	34	8	34	8	35	33	27.20	13.64
$C0$	27	30	24	31	32	28	33	28	25	37	29.50	3.92
$R0$	16	0	0	2	0	12	0	2	41	1	7.38	13.09
$D0$	0	0	0	0	1	0	0	0	0	0	0.12	0.32
SD	0.28	0.43	0.24	0.30	0.43	0.58	0.35	0.29	0.51	0.22	0.36	0.12
m_{sd}	28	1	20	17	8	14	11	20	8	1	13	9

Table 3: Optimized inputs for Germany, Argentina, Chile, Spain, Italy, South Korea, Portugal, Switzerland, Thailand and United States, respectively, for April 07, 2020. μ stands for the mean across countries, and STD for the standard deviation.

Inp.	GER	ARG	CHI	SPA	ITA	KOR	PO1	PO2	SWI	THA	USA	μ	STD
α	0.03	0.03	0.03	0.04	0.03	0.09	0.06	0.05	0.05	0.07	0.02	0.05	0.02
β	0.59	0.59	0.43	0.83	0.69	0.77	0.72	0.70	0.59	0.60	0.66	0.66	0.11
γ	2.43	2.20	2.21	2.61	3.14	2.12	2.37	2.37	3.02	1.63	2.18	2.39	0.44
δ	0.14	0.09	0.09	0.16	0.17	0.09	0.17	0.17	0.14	0.11	0.14	0.13	0.03
ζ	0.26	0.25	0.26	0.26	0.28	0.28	0.26	0.26	0.25	0.27	0.26	0.26	0.01
ε	0.09	0.07	0.11	0.10	0.12	0.10	0.11	0.11	0.10	0.06	0.10	0.10	0.02
m	0.76	0.77	0.75	0.73	0.72	0.81	0.76	0.76	0.75	0.76	0.78	0.76	0.03
l	0.013	0.007	0.008	0.017	0.032	0.000	0.018	0.018	0.016	0.008	0.015	0.014	0.009
rl	0.08	0.10	0.12	0.09	0.10	0.07	0.08	0.08	0.10	0.12	0.09	0.10	0.02
t_l	25	31	29	10	3	23	15	15	9	26	19	18	9
dur	26	14	23	26	28	24	29	29	17	22	24	24	5
c	0.26	0.17	0.15	0.37	0.38	0.15	0.34	0.34	0.36	0.24	0.33	0.28	0.09
f	0.47	0.44	0.46	0.42	0.55	0.42	0.47	0.47	0.47	0.44	0.45	0.46	0.04
$E0$	19	20	21	20	20	20	20	20	20	20	20	20	0
$I0$	39	41	42	40	41	41	40	40	41	41	41	41	1
$H0$	7	8	8	8	8	8	8	8	8	8	8	8	0
$C0$	27	28	29	28	28	28	28	28	28	28	28	28	0
$R0$	16	0	0	2	1	12	0	0	2	35	3	5	11
$D0$	0	0	0	0	2	0	0	0	0	1	0	0	1
SD	0.36	0.38	0.39	0.24	0.40	0.57	0.47	0.47	0.39	0.50	0.21	0.40	0.11
m_{sd}	27	3	19	14	14	14	27	30	24	10	8	16	9

Table 4: Inverse of the model optimized coefficients of γ, δ, ζ, ε representing latent, infectious, hospitalization and critical cases mean duration in days, as well as the model estimated basic reproductive number (**R0)** for April 14, 2020 for Germany, Argentina, Chile, Spain, Italy, South Korea, Portugal, Switzerland, Thailand and United States, respectively. μ stands for the mean across countries, and STD for the standard deviation.

Country	Latent	Infectious	Hospitalized	Critical	**R0**
GER	0.5	8.8	3.9	9.9	2.89
ARG	0.3	11.3	3.6	12.6	4.77
CHILE	0.6	11.4	3.9	13.4	2.28
SPAIN	0.3	7.0	3.4	10.0	3.10
ITALY	0.3	5.9	3.6	8.5	3.25
KOREA	0.5	11.3	3.8	10.1	2.68
PORT	0.4	7.7	3.7	9.2	2.86
SWITZ	0.4	6.8	4.1	9.6	2.09
THAI	0.5	9.2	3.9	16.6	2.39
USA	0.4	7.6	3.6	8.8	3.78
μ	0.4	8.7	3.7	10.9	3.0
STD	0.1	2.1	0.2	2.6	0.8

Table 5: Model results for total number of infected, hospitalized, ICU patients, recovered and deaths, as well as peak day hospitalization, peak day ICU, number of deaths with no treatment (D_OH) and percentage of death caused by lack of treatment for Germany, Argentina, Chile, Spain, Italy, South Korea, Portugal, Switzerland, Thailand and United States, respectively. μ stands for the mean across countries, and STD for the standard deviation.

Country	Infect	Hosp	peak_Hos	ICU	peak_ICU	REC	Deaths	D_OH	DOH/D
GER	250292	75590	7492	17753	3856	239655	10671	2380	0.22
ARG	200948	44095	2741	6176	1178	197447	3569	1070	0.30
CHI	20231	4834	484	939	249	19804	492	96	0.20
SPA	1395656	438540	35317	148866	30586	1309648	86074	19090	0.22
ITA	616554	205766	16044	78527	13280	550169	66452	23262	0.35
KOR	24153	5103	552	787	179	23863	325	0	0.00
POR	87655	28132	2246	9343	1683	82295	5427	1207	0.22
SWI	52202	16509	1905	4729	1077	49365	2873	611	0.21
THA	4647	1323	127	321	93	4592	115	0	0.00
USA	2634990	907921	83621	360008	71095	2437866	197194	32934	0.17
μ	528733	172781	15053	62745	12328	491470	37319	8065	0.19
STD	855702	292350	26466	115002	22772	792481	64072	12240	0.11

Table 6: Seven days percentage errors and absolute percentage errors in estimations for the active cases (C) and cumulated deaths (D), considering the model run for April 7 and comparing to real data for April 14 for Germany, Argentina, Chile, Spain, Italy, South Korea, Portugal, Switzerland, Thailand and United States, respectively. μ stands for the mean across countries, and STD for the standard deviation.

Country	Cases	Deaths	Abs C	Abs D
GER	-10	40	10	40
ARG	19	23	19	23
CHI	15	42	15	42
SPA	03	10	03	10
ITA	06	27	06	27
KOR	14	06	14	06
POR	26	31	26	31
SWI	-16	14	16	14
THA	-16	27	16	27
USA	-06	33	06	33
GER	30	50	30	50
μ	06	28	15	28
STD	16	14	08	14

Figure 1

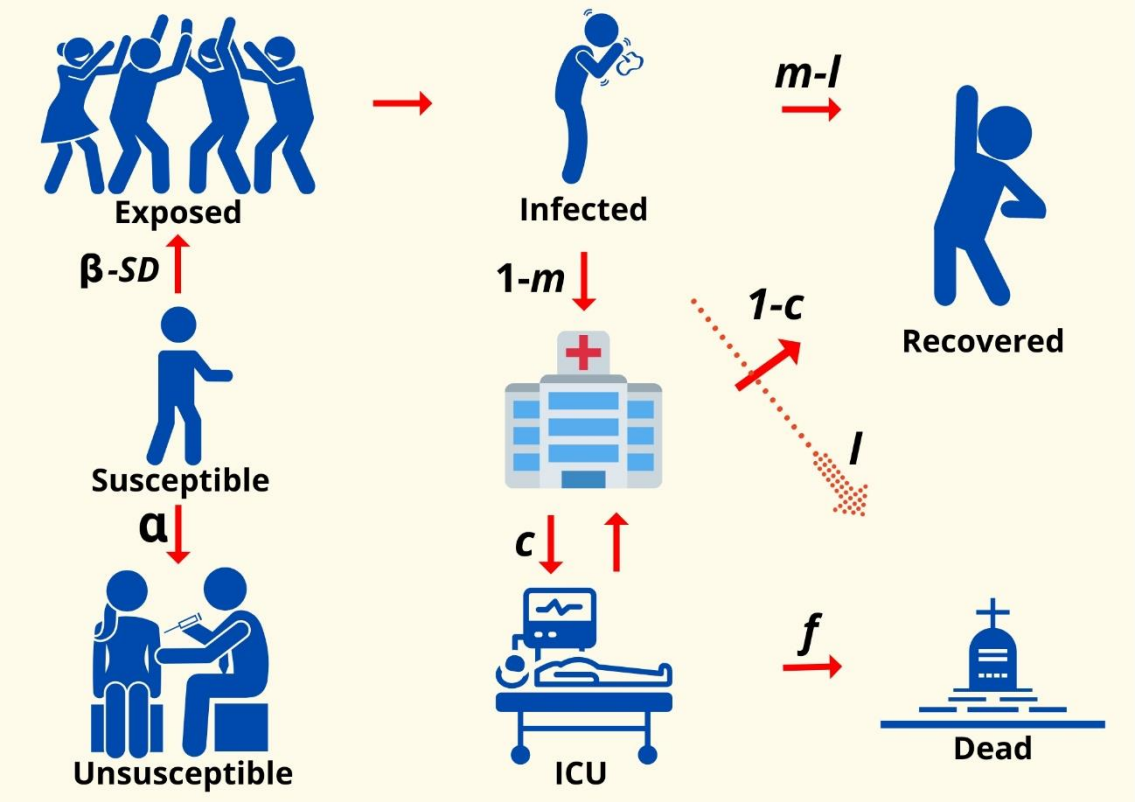

Exposed
Infected
m-I
Recovered
β-SD
1-m
1-c
Susceptible
I
α
c
f
Dead
Unsusceptible
ICU

Figure 1 Black and White

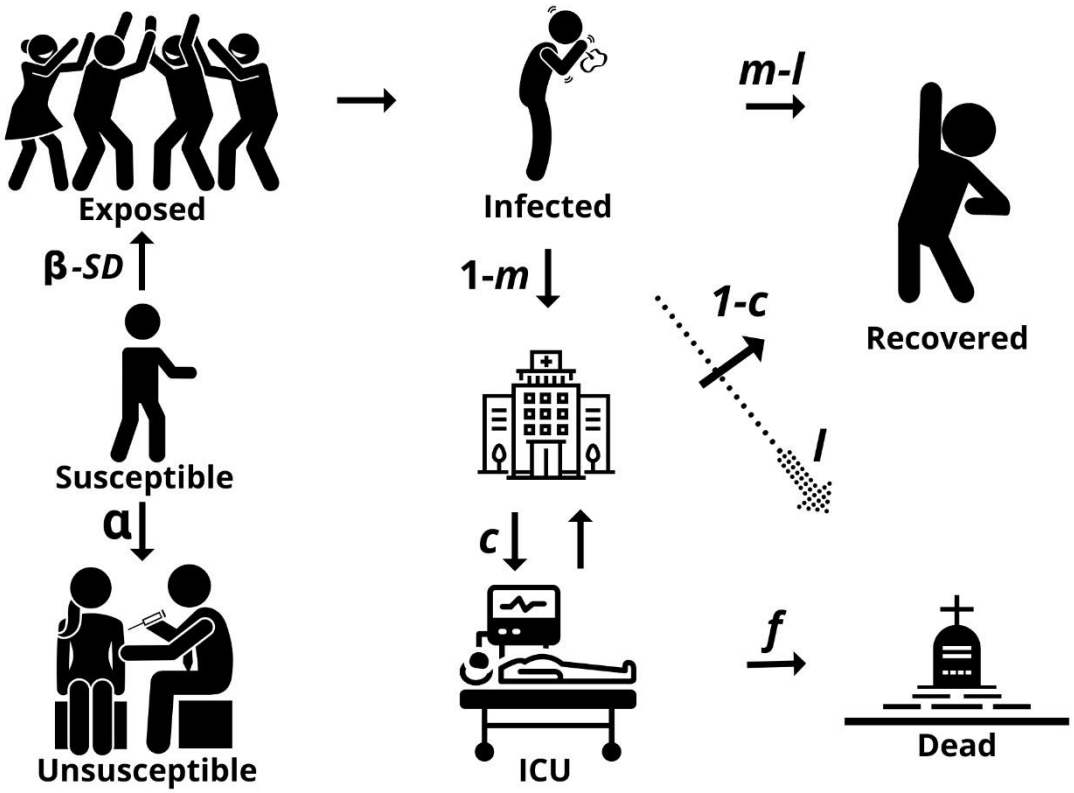
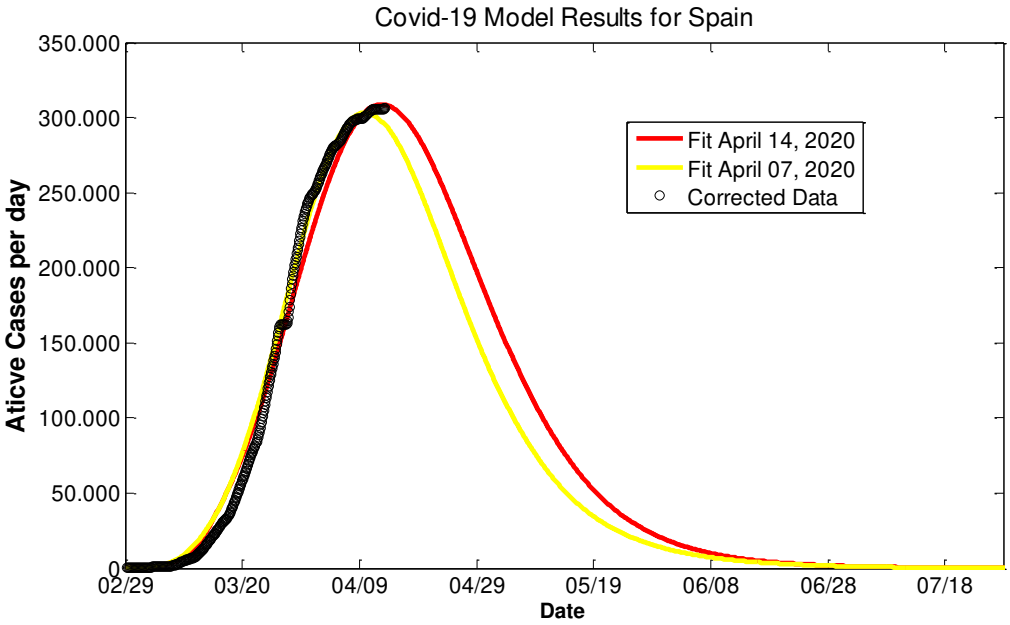


Figure 2


Covid-19 Model Results for Spain
Aticve Cases per day
350.000
300.000
250.000
200.000
150.000
100.000
50.000
0
Fit April 14, 2020
Fit April 07, 2020
Corrected Data
02/29
03/20
04/09
04/29
05/19
06/08
06/28
07/18
Date

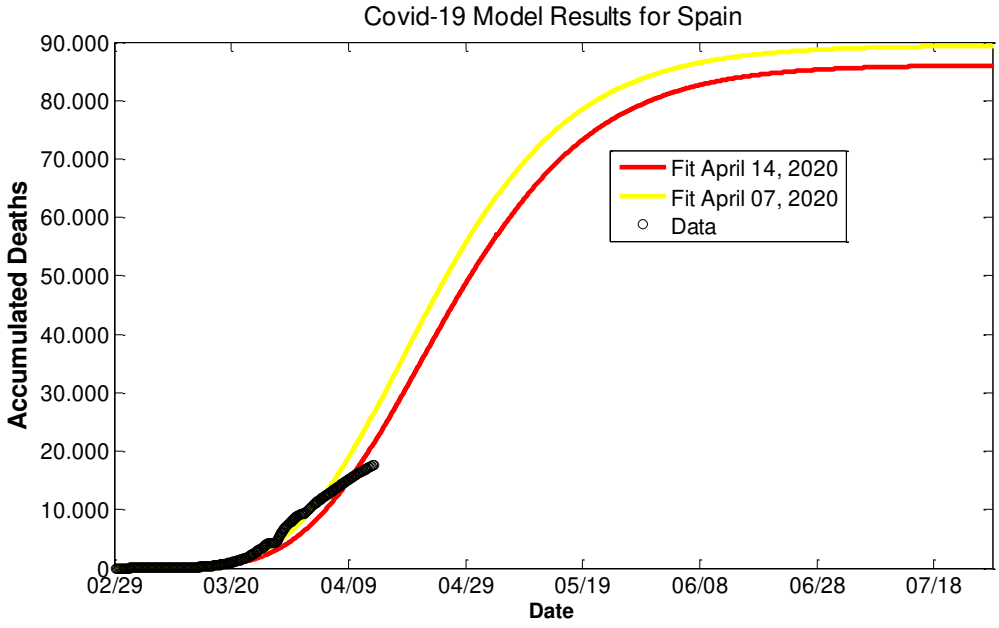

Covid-19 Model Results for Spain
Accumulated Deaths
90.000
80.000
70.000
60.000
50.000
40.000
30.000
20.000
10.000
0
Fit April 14, 2020
Fit April 07, 2020
Data
02/29
03/20
04/09
04/29
05/19
06/08
06/28
07/18
Date

Figure 2 Black and White

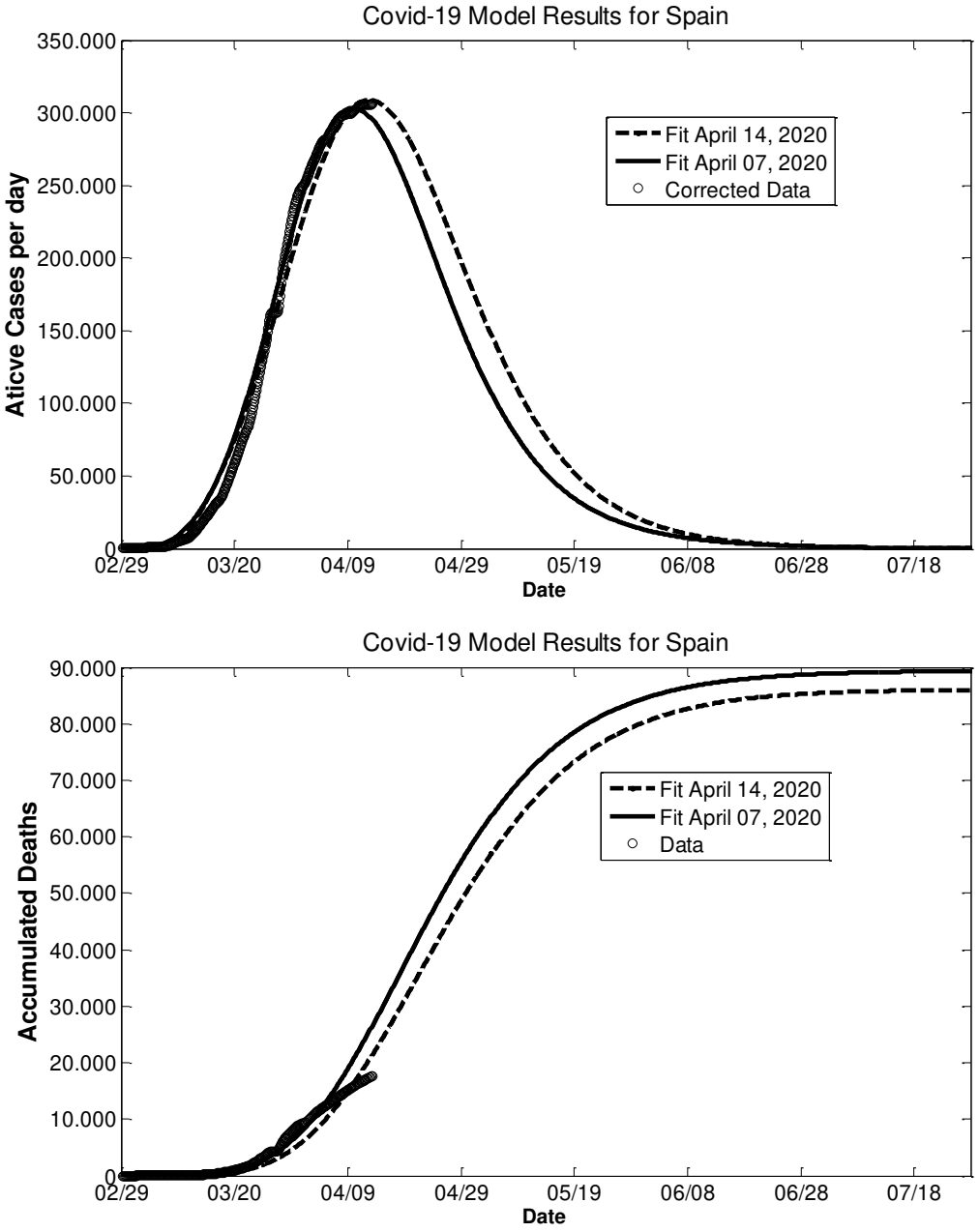


Figure 3

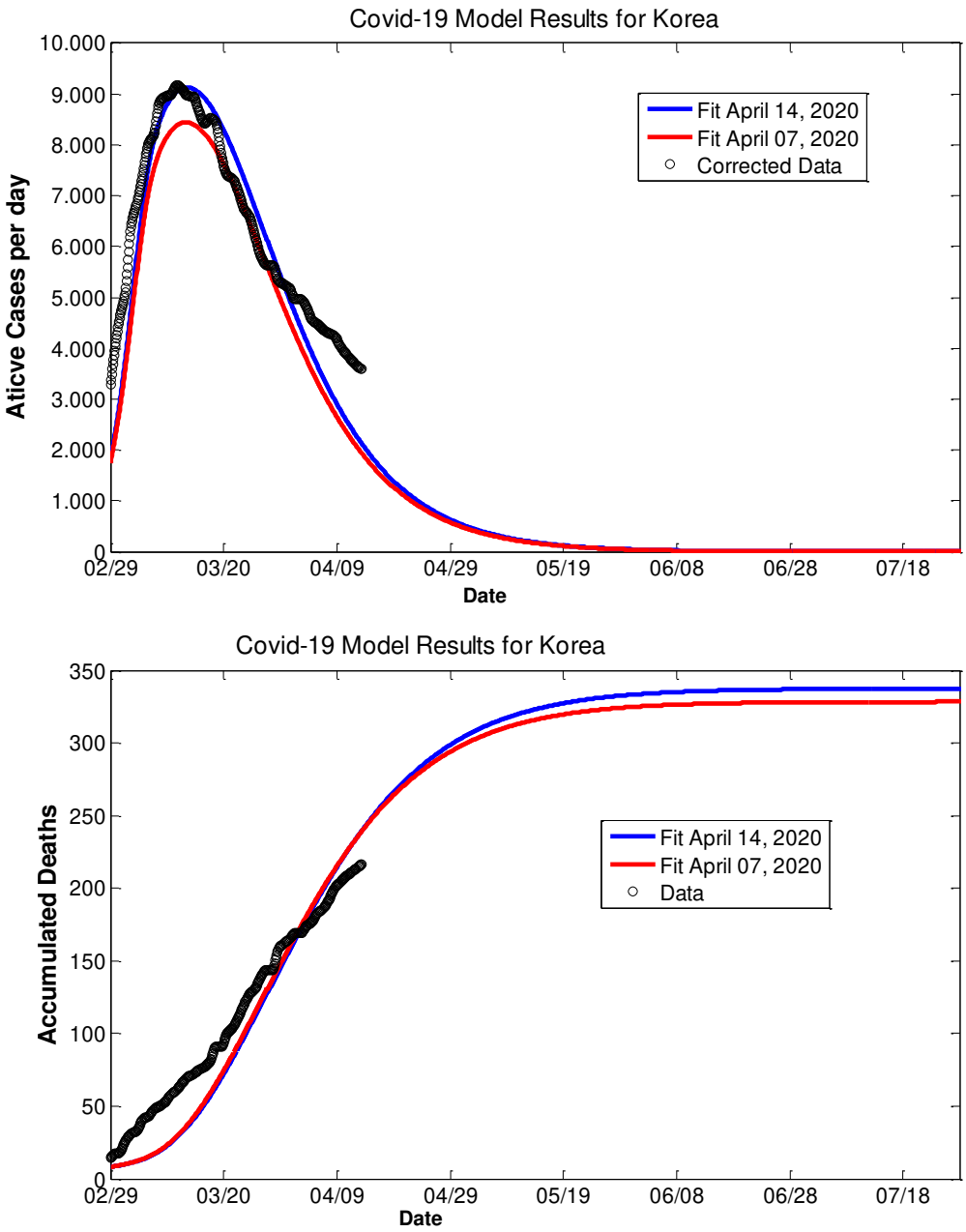

Covid-19 Model Results for Korea
Aticve Cases per day
10.000
9.000
8.000
7.000
6.000
5.000
4.000
3.000
2.000
1.000
0
02/29
03/20
04/09
04/29
05/19
06/08
06/28
07/18
Date
Fit April 14, 2020
Fit April 07, 2020
Corrected Data
Covid-19 Model Results for Korea
Accumulated Deaths
350
300
250
200
150
100
50
0
02/29
03/20
04/09
04/29
05/19
06/08
06/28
07/18
Date
Fit April 14, 2020
Fit April 07, 2020
Data

Figure 3 Black and White

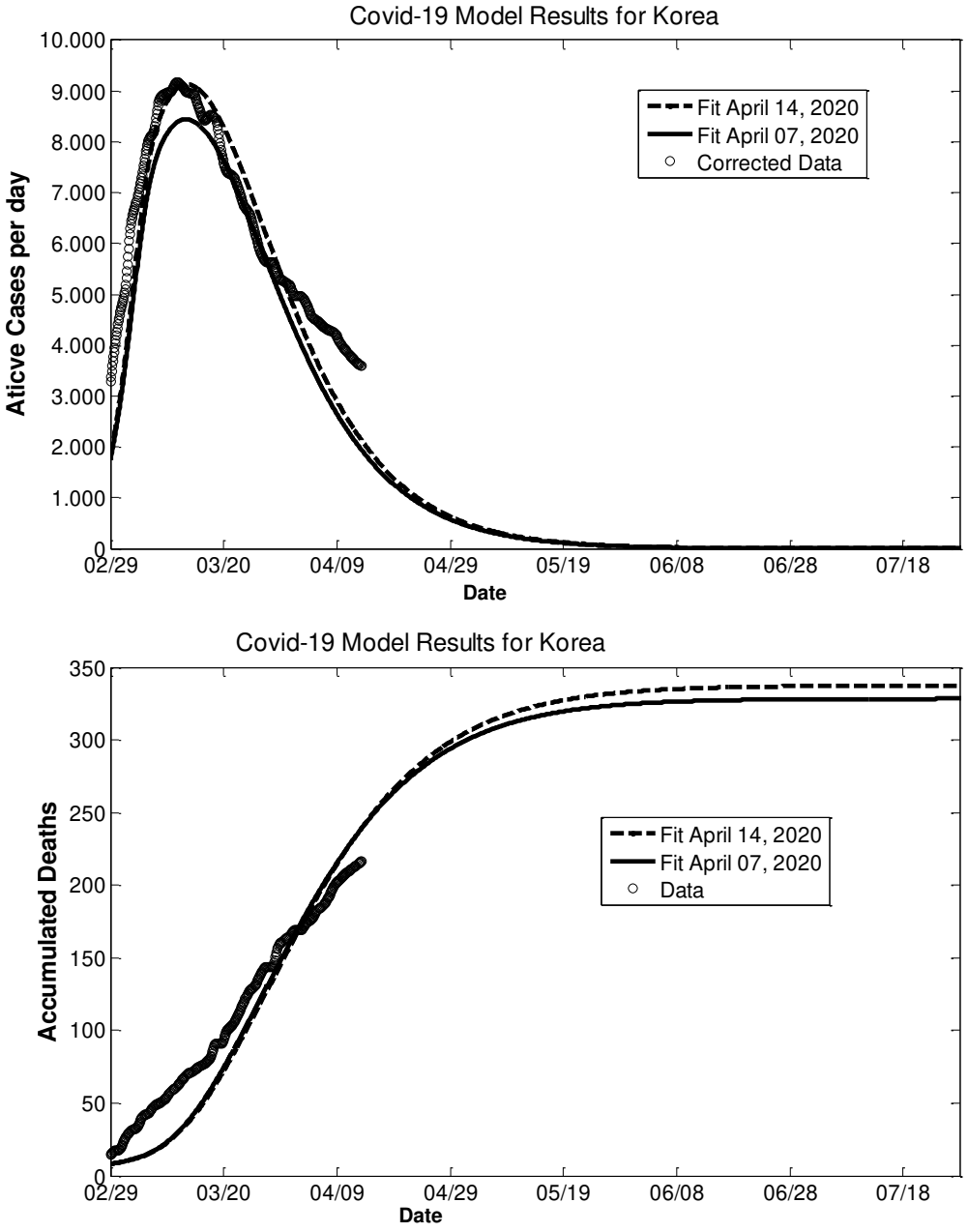


Figure 4

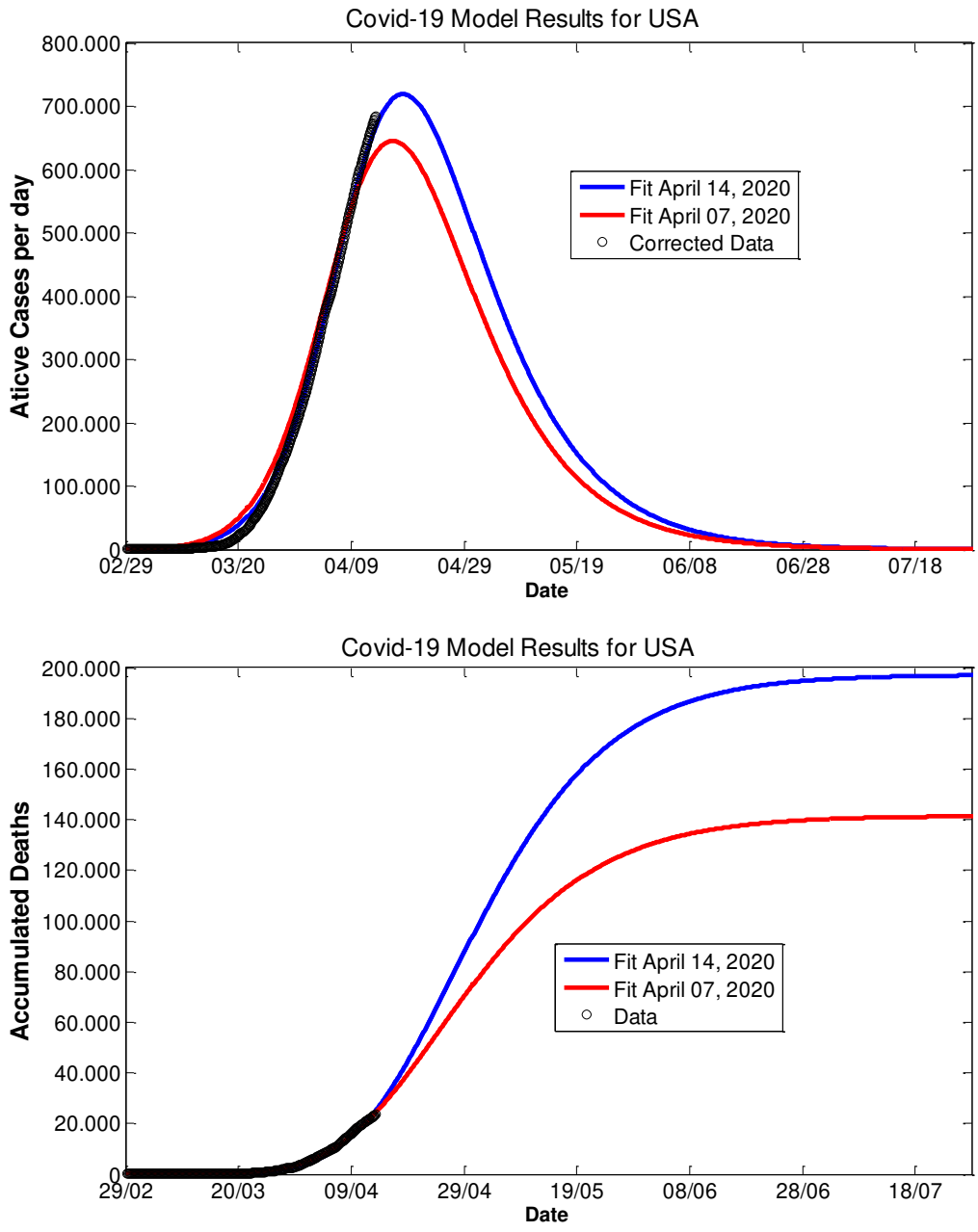

Covid-19 Model Results for USA
Aticve Cases per day
800.000
700.000
600.000
500.000
400.000
300.000
200.000
100.000
0
02/29
03/20
04/09
04/29
05/19
06/08
06/28
07/18
Date
Fit April 14, 2020
Fit April 07, 2020
Corrected Data
Covid-19 Model Results for USA
Accumulated Deaths
200.000
180.000
160.000
140.000
120.000
100.000
80.000
60.000
40.000
20.000
0
29/02
20/03
09/04
29/04
19/05
08/06
28/06
18/07
Date
Fit April 14, 2020
Fit April 07, 2020
Data

Figure 4 Black and White

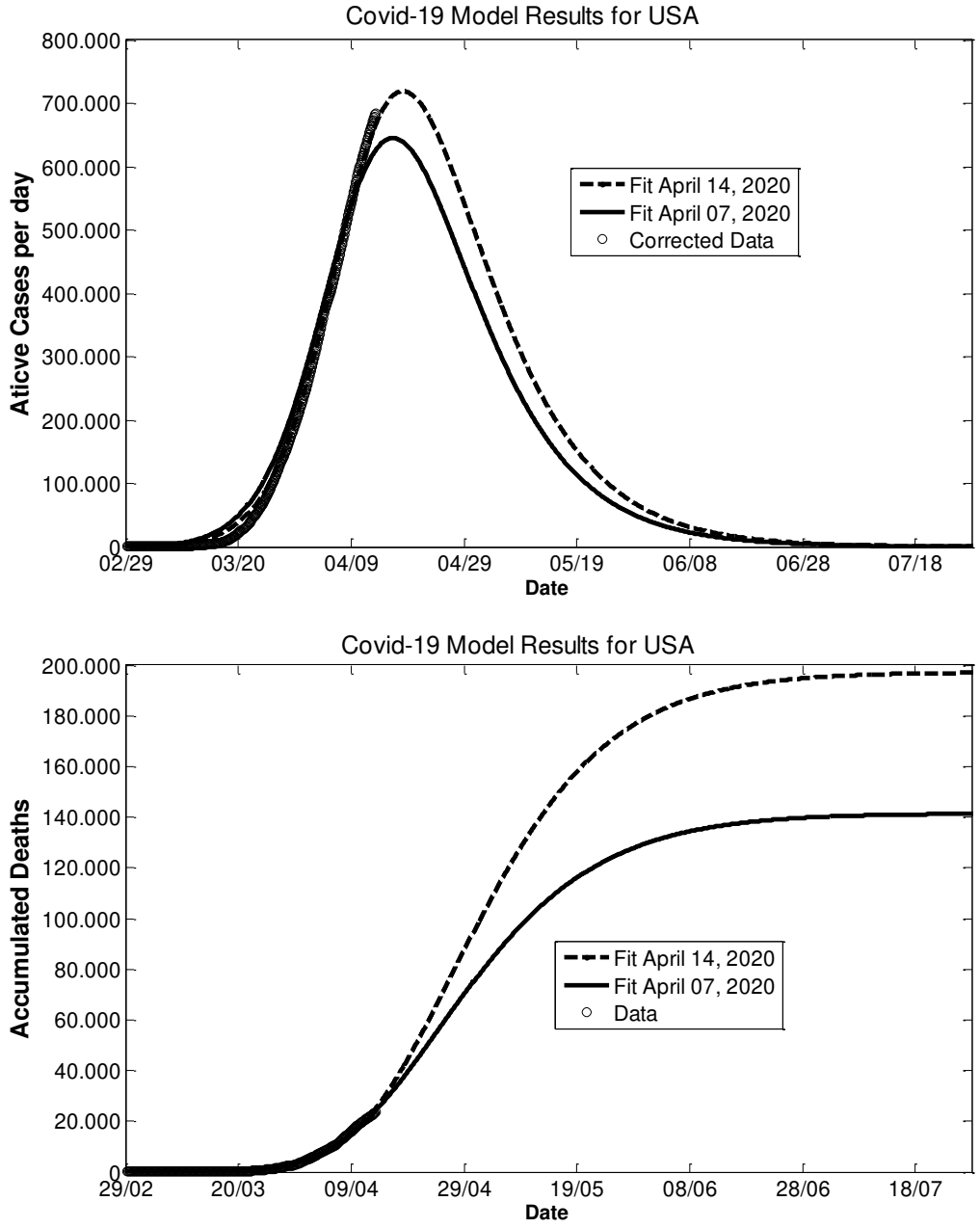


Figure 5

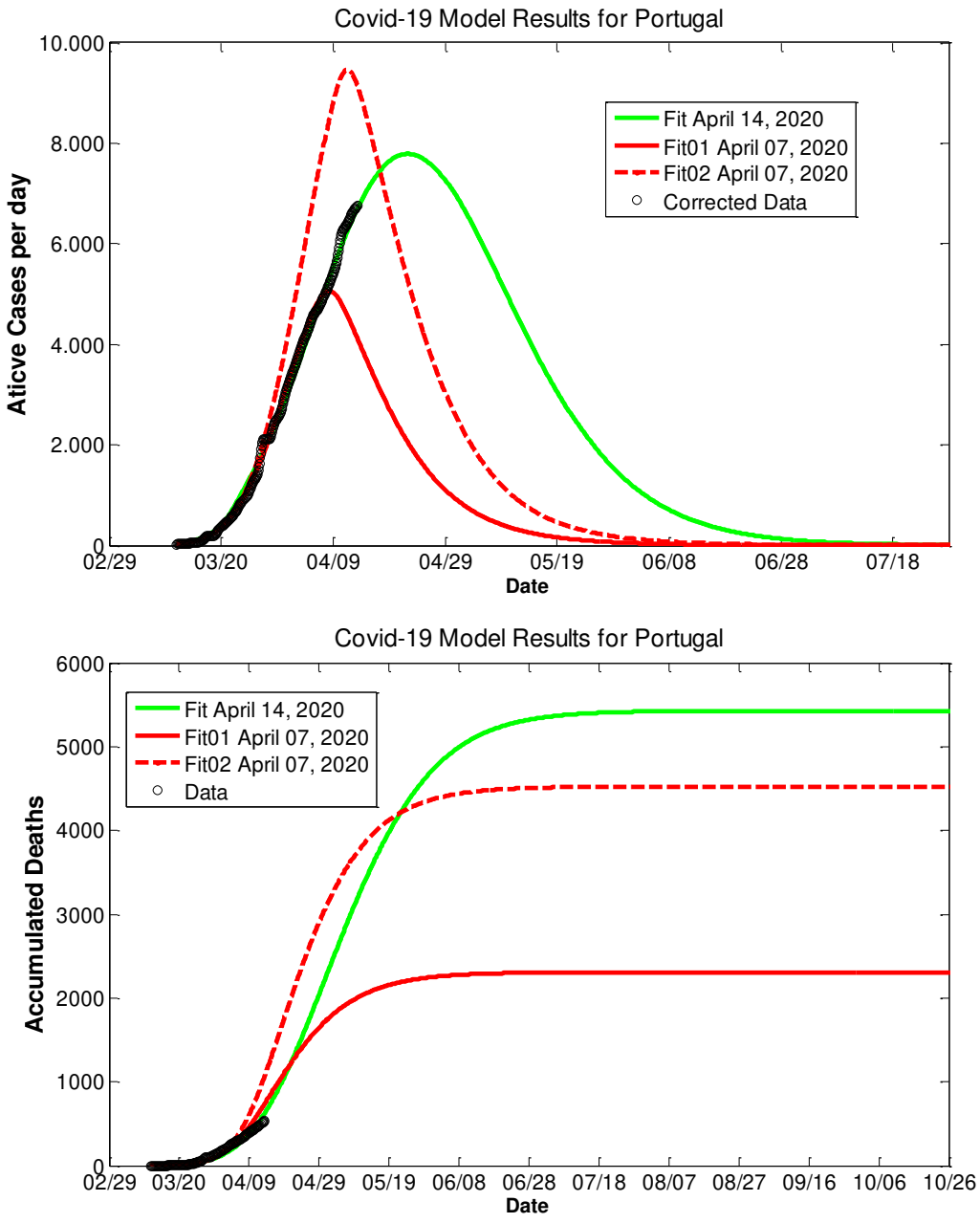
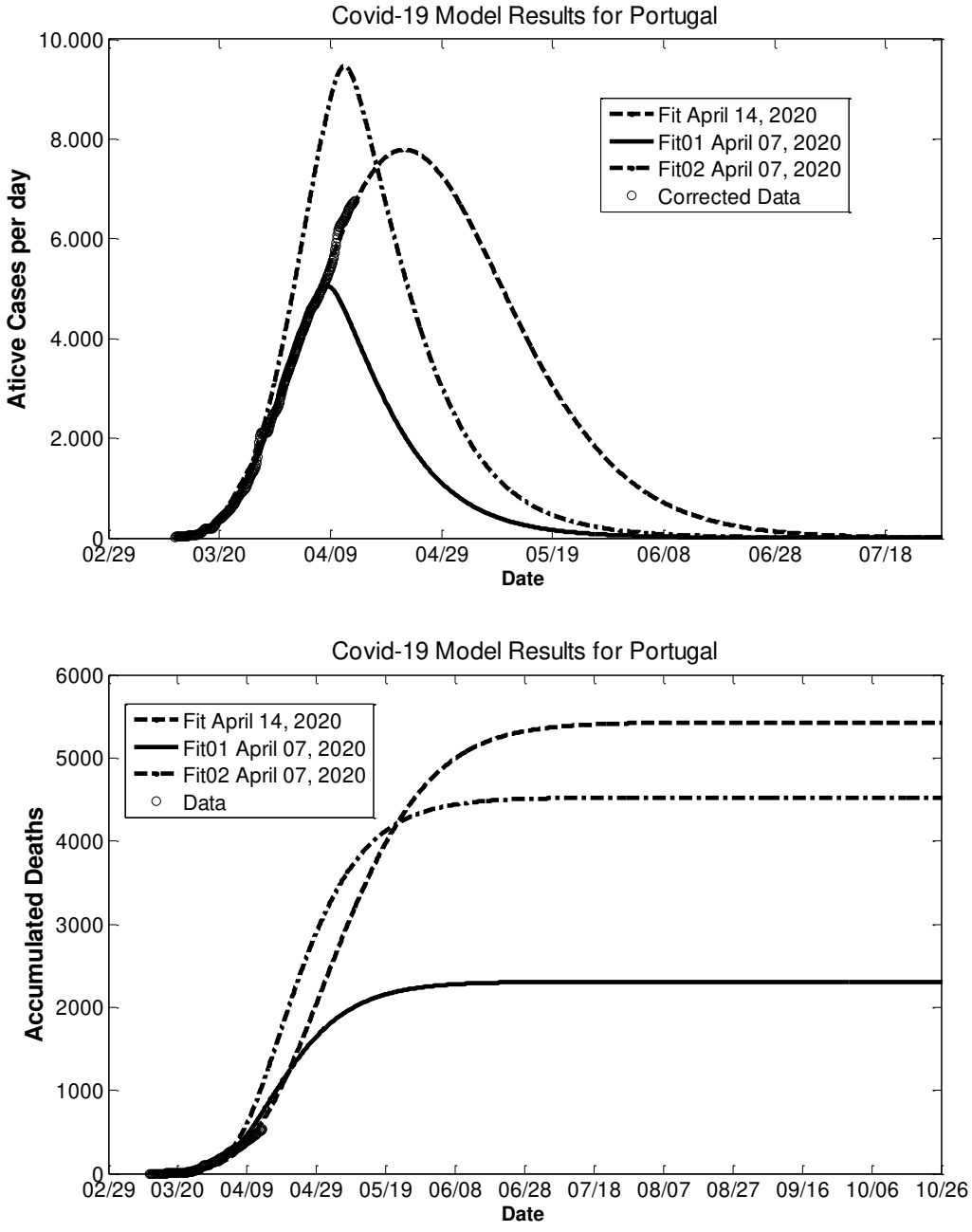

Covid-19 Model Results for Portugal
Aticve Cases per day
Date
Fit April 14, 2020
Fit01 April 07, 2020
Fit02 April 07, 2020
Corrected Data
10.000
8.000
6.000
4.000
2.000
0
02/29
03/20
04/09
04/29
05/19
06/08
06/28
07/18
Covid-19 Model Results for Portugal
Accumulated Deaths
Date
Fit April 14, 2020
Fit01 April 07, 2020
Fit02 April 07, 2020
Data
6000
5000
4000
3000
2000
1000
0
02/29
03/20
04/09
04/29
05/19
06/08
06/28
07/18
08/07
08/27
09/16
10/06
10/26

Figure 5 Black and White



Figures

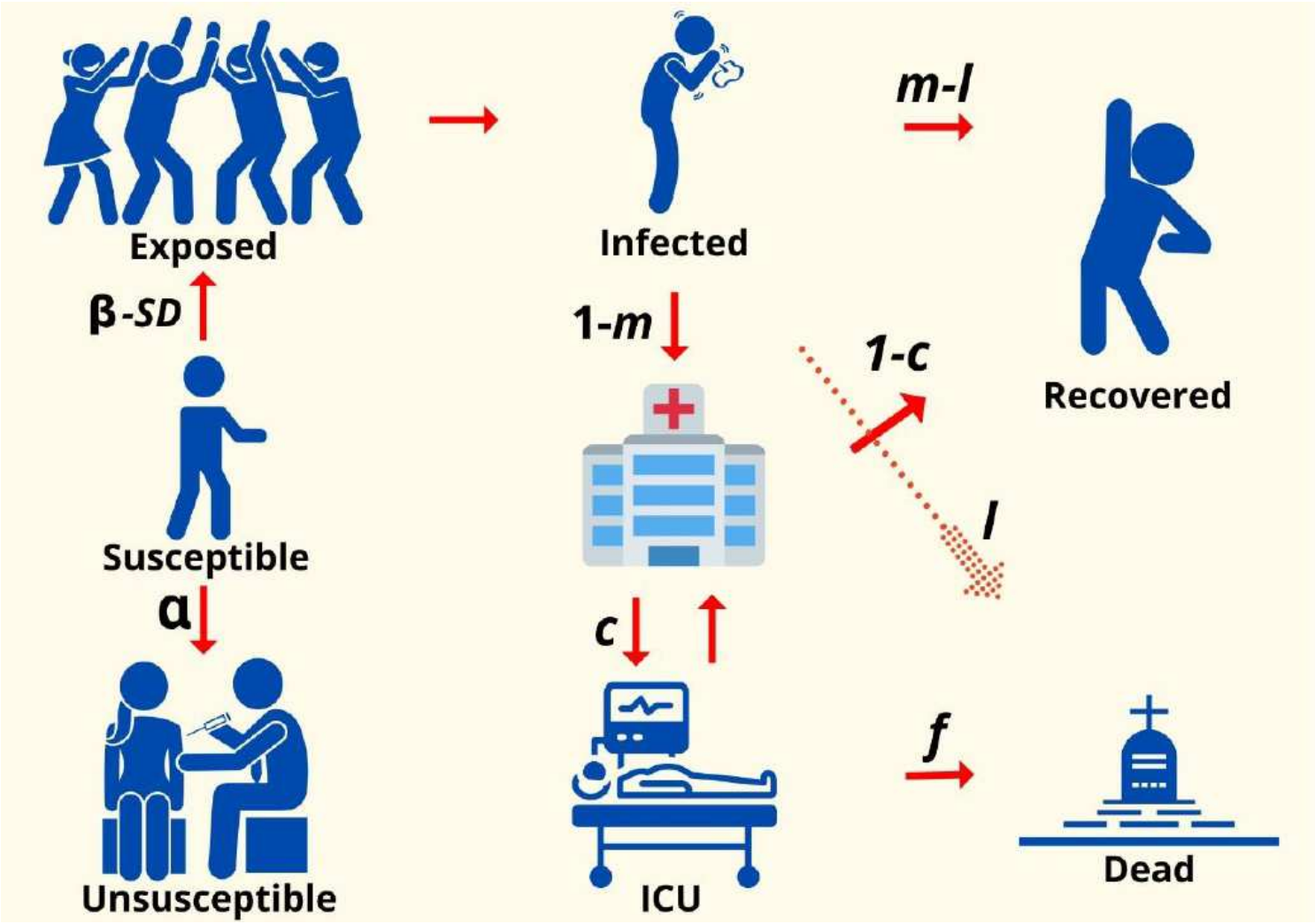


Figure 1

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.

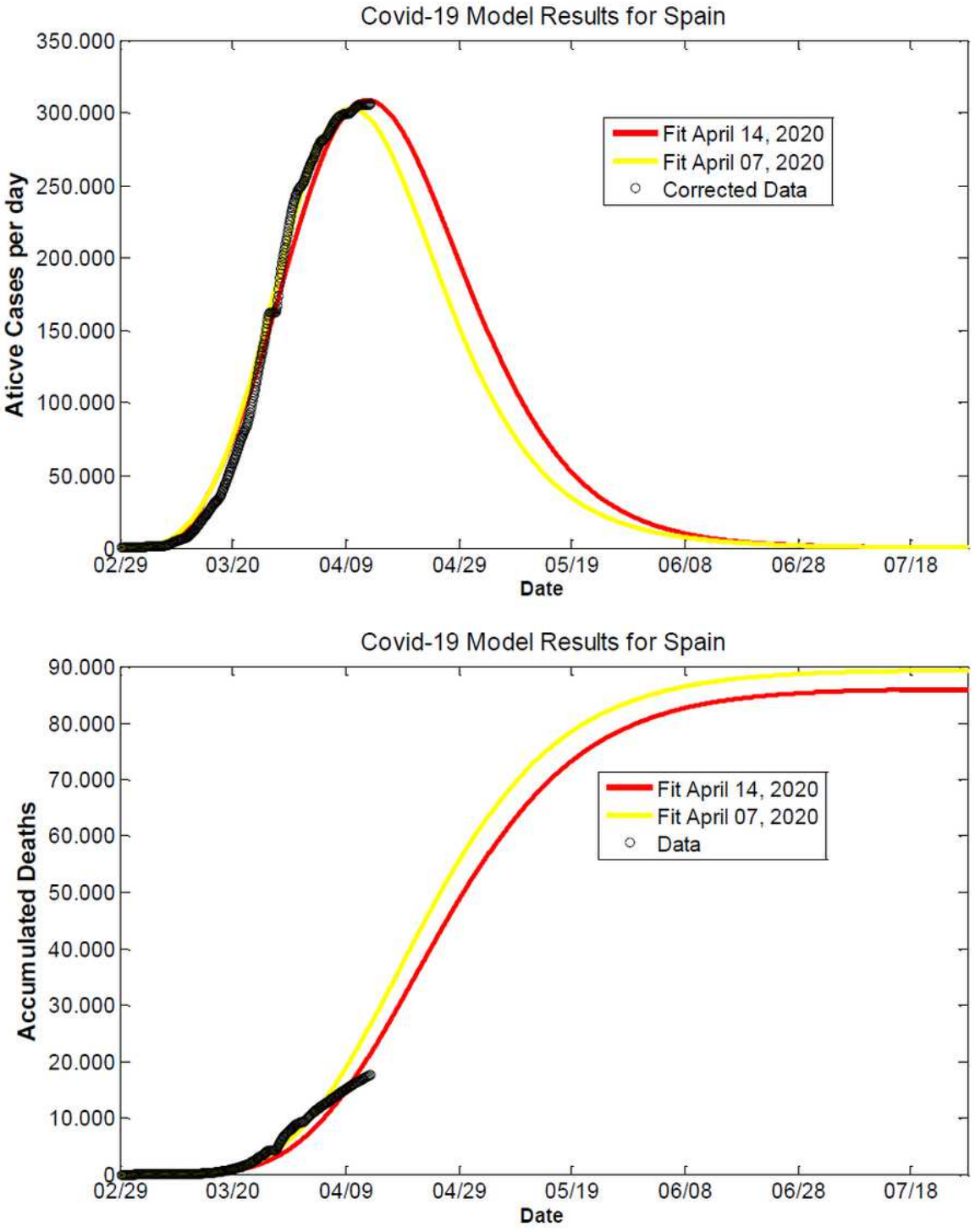


Figure 2

Model results for active cases (a) and accumulated deaths (b) for Spain, using data from April 07, 2020 and April 14, 2020.

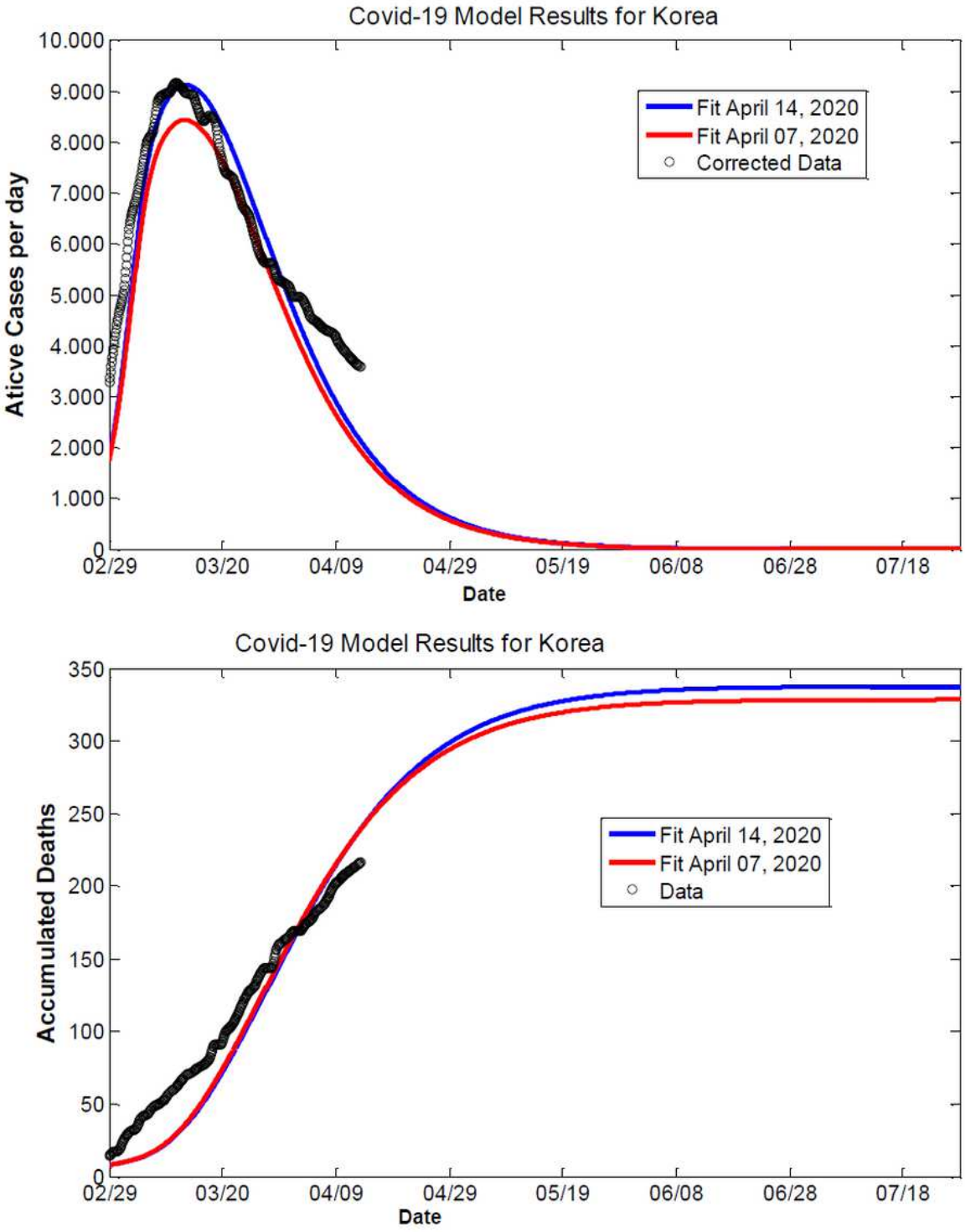


Figure 3

Model results for active cases (a) and accumulated deaths (b) for South Korea, using data from April 07, 2020 and April 14, 2020.

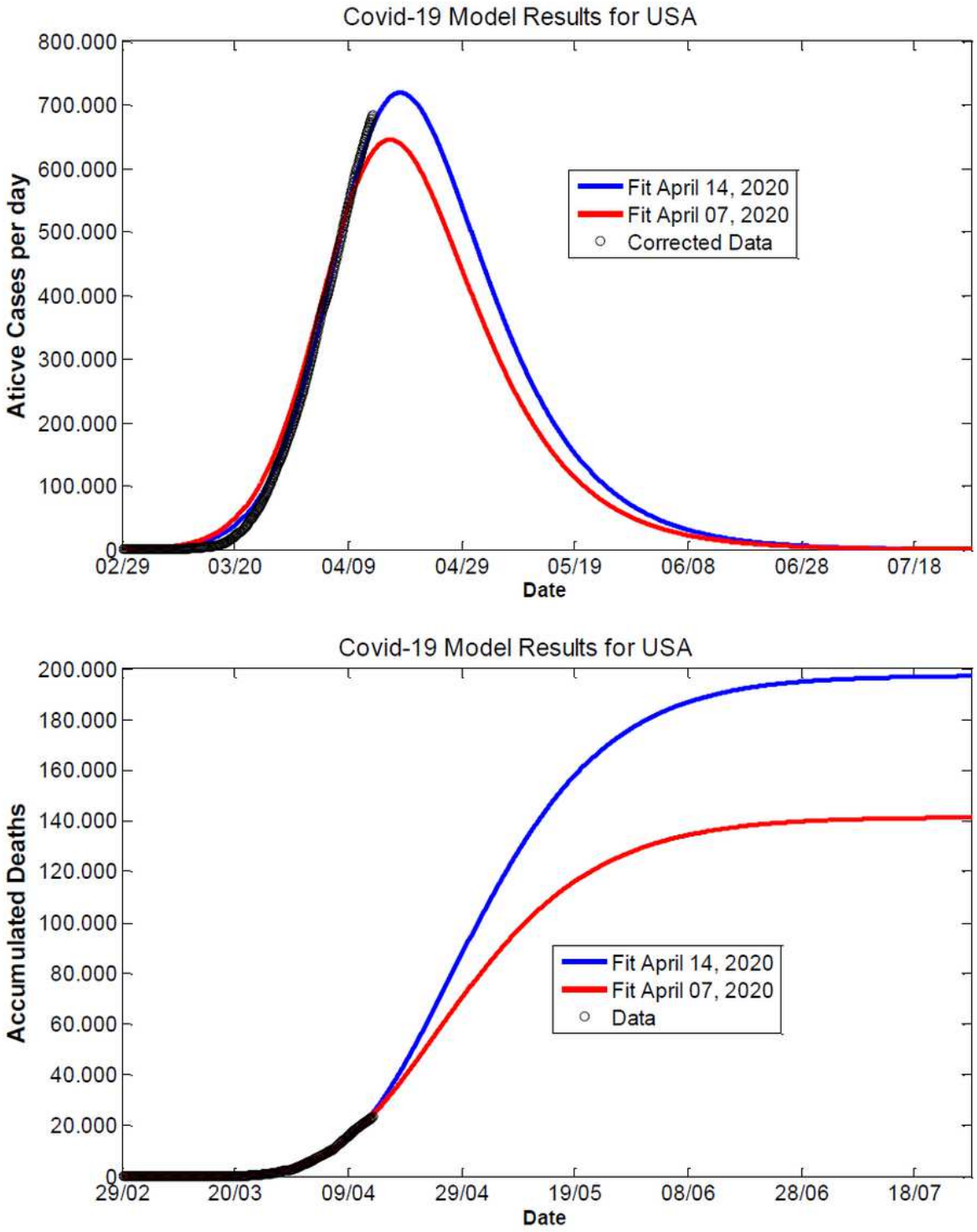


Figure 4

Model results for active cases (a) and accumulated deaths (b) for United States, using data from April 07, 2020 and April 14, 2020.

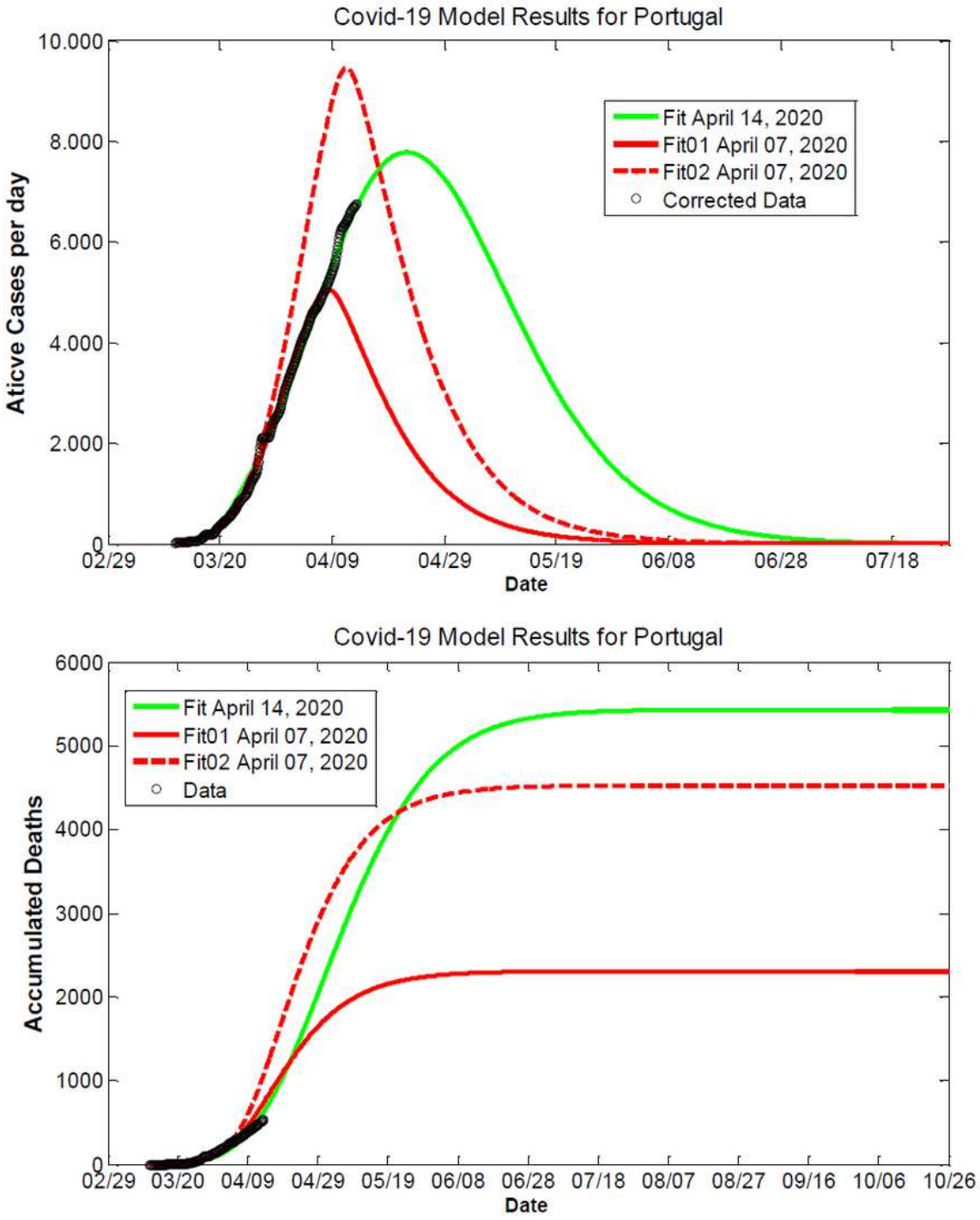


Figure 5

Model results for active cases (a) and accumulated deaths (b) for Portugal, using data from April 07, 2020 and April 14, 2020. Note that there are two possible optimizations for Portugal using data from April 07, 2020.