Spreading of COVID-19 in Brazil: Impacts and uncertainties in social distancing strategies

Diego T. Volpatto[a,c,*], Anna Claudia M. Resende[a], Lucas dos Anjos[a], João V. O. Silva[a], Claudia M. Dias[b], Regina C. Almeida[a], Sandra M. C. Malta[a]

[a] *LNCC, Laboratório Nacional de Computação Científica, Petrópolis, RJ, Brazil*
[b] *UFRRJ, Universidade Federal Rural do Rio de Janeiro, Nova Iguaçu, RJ, Brazil*
[c] *ESSS, Engineering Simulation and Scientific Software, Florianópolis, SC, Brazil*

Abstract

Brazil's continental dimension poses a challenge to the control of the spread of COVID-19. Due to the country specific scenario of high social and demographic heterogeneity, combined with limited testing capacity, lack of reliable data, under-reporting of cases, and restricted testing policy, the focus of this study is twofold: (i) to develop a generalized SEIRD model that implicitly takes into account the quarantine measures, and (ii) to estimate the response of the COVID-19 spread dynamics to perturbations/uncertainties. By investigating the projections of cumulative numbers of confirmed and death cases, as well as the effective reproduction number, we show that the model parameter related to social distancing measures is one of the most influential along all stages of the disease spread and the most influential after the infection peak. Due to such importance in the outcomes, different relaxation strategies of social distancing measures are investigated in order to determine which strategies are viable and less hazardous to the population. The results highlight the need of keeping social distancing policies to control the disease spread. Specifically, the considered scenario of abrupt social distancing relaxation implemented after the occurrence of the peak of positively diagnosed cases can prolong the epidemic, with a significant increase of the projected numbers of confirmed and death cases. An even worse scenario could occur if the quarantine relaxation policy is implemented before evidence of the epidemiological control, indicating the importance of the proper choice of when to start relaxing social distancing measures.

Keywords: COVID-19, Generalized SEIRD Model, Social Distancing, Reproduction Number

Introduction

At the end of 2019, the world was taken by the news about the outbreak in China of a new Coronavirus called SARS-CoV-2, which stands for *Severe Acute Respiratory Syndrome Coronavirus 2*. The associated disease, called COVID-19 by the World Health Organization (WHO), rapidly spread around the world and is currently considered a pandemic. As of May 26, there have been 216 countries and territories with confirmed cases, the majority of them with local transmission [1]. Coronavirus belongs to a group of viruses that are common in humans and are responsible for 15% to 30% of cases of common cold [2]. The SARS-CoV-2 has spread rapidly and has already affected

[‡] DTV, ACMR, and JVOS performed computer simulations. All authors contributed to the conceptualization of this study, methodology, writing, and approval of the final version of the manuscript.

[‡‡] The authors declare no competing interest.

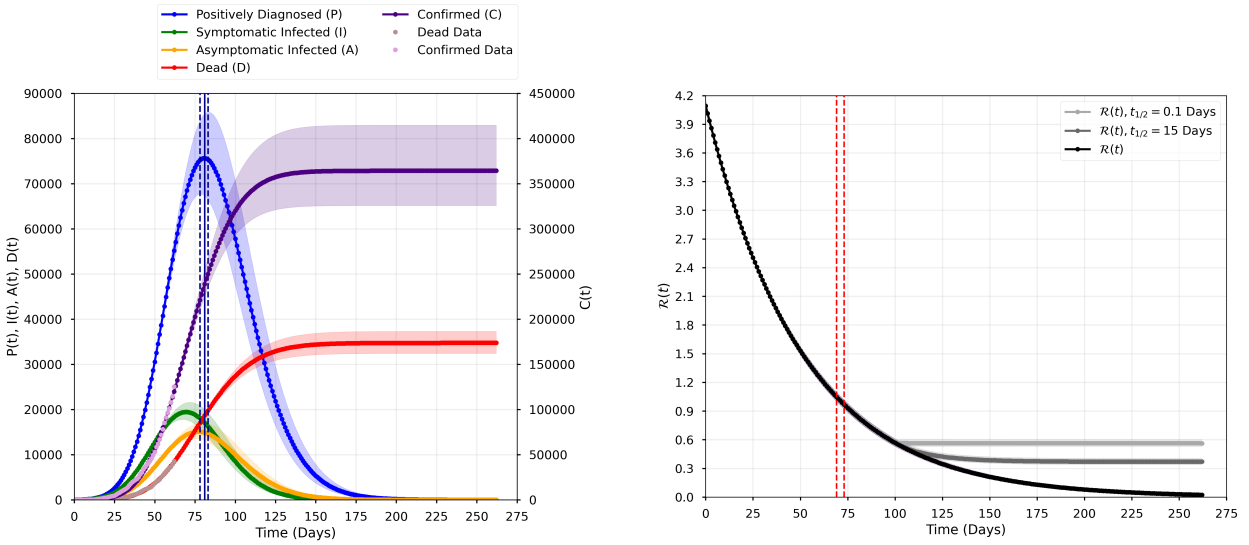
[*] Corresponding Author. E-mail: volpatto@lncc.br

a significant part of the world's population [3]. The within-host dynamics seem to be very heterogeneous and the severity of the disease varies widely. Based on currently available information, elderly people and people of any age with comorbidities, such as hypertension, cardiovascular disease, and diabetes, are among those at most risk of severe illness from COVID-19. In most locations, the burden of COVID-19 cases requiring medical intensive care exceeds the available medical resources, worsening the situation. A great effort has been made worldwide to better understand the disease and possible mechanisms to control it. Measures such as widespread testing, and reducing social contacts for some age segments, or for the wider population, are known to decrease the speed of the disease spread and the case fatality [4]. However, countries often have to deal with limit testing capabilities for COVID-19 which increases the uncertainty associated with how the disease behaves and is likely to complicate the management of the policies to mitigate and control it. One form to deal with this lack of information is through mathematical modeling, using diversified modelling techniques (see [5] for a review). Population-based models that split individuals into classes are widely used, and is the approach followed in this work.

WHO has already recognized that testing for COVID-19 is a key way to know how the virus spread and to provide insights on how to respond to it, although widespread testing is low in most countries in the world. Current Brazilian policy is to test only severely ill and healthcare practitioners people. Even so, there is great uncertainty in the released data. Sub-notification of infected cases is a major problem: BR currently conducts around 4,000 tests per million of the population [3], one of the lowest rates in the world. Likewise, the number of deaths may be underestimated for the same reason. According to [6], only 8% of the infected cases are reported/diagnosed. Overall, one of the main difficulties in developing predictive mathematical models of the COVID-19 in Brazilian localities is the lack of reliable data to support parameter choices [7, 8]. In the face of the current scenario, we developed a seven-compartment model that implicitly considers the social distancing policy that isolates individuals from the infection for a period of time. We focus on understanding the model response to perturbations/uncertainties. We perform sensitivity analysis and quantify the uncertainties in model outcomes, mainly the cumulative numbers of confirmed and death cases, and the effective reproduction number. We also investigate how to obtain a desired effect/control via a modification to the system subject to uncertainties. We first apply the model to the whole BR scenario and then investigate its behavior for the state of Rio de Janeiro (RJ), which was one of the first states in BR to recommend and implement social distancing policy, beginning on March 17, 2020 [9].

Results

We consider a generalized SEIRD mathematical model containing the following epidemiological compartments: Susceptible (S), Exposed (E), Asymptomatic Infected (A), Symptomatic Infected (I), Recovered or Removed (R), Positively Diagnosed individuals (P) [10], and Cumulative Dead (D). There is also a variable that is not part of the dynamical system and corresponds to the cumulative number of Confirmed Cases (C). Due to the current Brazilian policy only severely ill and healthcare practitioners people are tested. Therefore, it is important a separation in P individuals most likely to need hospitalization, in contrast to a limited number of hospital beds. Moreover, a proportion of S, E, A, and I individuals are removed at a fixed rate to the R compartment, which corresponds to quarantine measures and recovered individuals, as proposed in [11]. This mathematical model is denoted SEAIRPD-Q, in which Q stands for the quarantine (Materials and Methods). In order to address the understanding of quarantine policies, in the next sections we analyze different scenarios related to quarantine measures through modifications in the aforementioned removal rate.



(a) Prediction of P, I, A, D, and C (right y-axis) in BR. The vertical line indicates the curve peak, and the dashed lines display the corresponding 95% CI.

(b) Evolution of $\mathcal{R}(t)$ in BR. Dashed lines indicate 95% CI around $\mathcal{R}(t) = 1$. The black curve refers to the simulation shown in Fig. 1a, and others to simulations with different distancing measures.

Figure 1: Dynamics of the COVID-19 in BR modeled with the available data (Materials and Methods). The lines indicate the simulation with the maximum *a posteriori* (MAP) estimates, with the shaded colors indicating the 95% credible interval (CI). Simulation days 0 and 81 corresponds to March 5, 2020 and May 24, 2020, respectively.

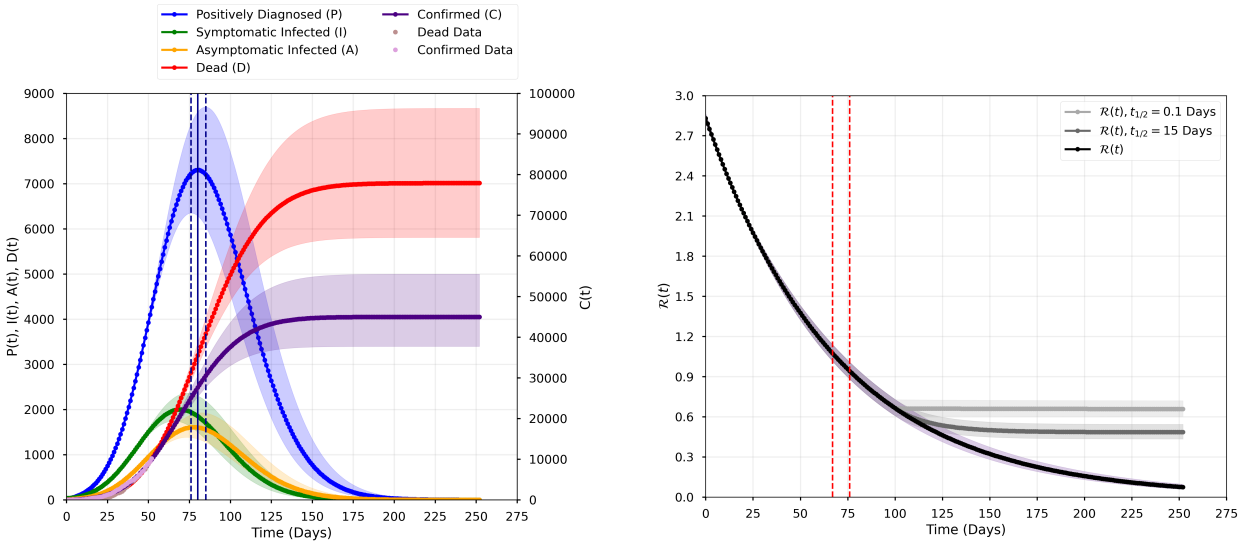
Current stage and trends for COVID-19 in BR and RJ

To properly perform projections and study the quarantine scenarios of interest, we fitted the SEAIRPD-Q model (Materials and Methods) to the data of BR and RJ (separately). We employed a Bayesian calibration of some model parameter values, with remaining parameter values gathered from the literature.

Considering the maintenance of current social distancing policies, the predictions for the COVID-19 pandemic in BR is showed in Fig. 1. According to the result of Fig. 1a, the peak of compartment P will occur on the simulation day 81 (95% CI: 78–83) with around 75.7 thousand (95% CI: 67.7–85.8) people simultaneously infected (denoted "active cases", hereinafter), and D is expected to be around 34.7 thousand (95% CI: 32.4–37.3) in the considered time range (Materials and Methods). According to Fig. 1a, C in the considered time range will be around 364.2 thousand (95% CI: 325.7–414.8).

The basic reproduction number, $\mathcal{R}_0$, is a parameter that indicates how fast the disease will spread in an epidemic early stage, caused by secondary infections after one individual is infected [12]. We calculated this quantity and the corresponding effective reproduction number, $\mathcal{R}(t)$ (SISAppendix A.3). The $\mathcal{R}_0$ calculated by means of MAP for BR is 4.09, and the distribution of $\mathcal{R}(t)$ along time is displayed in Fig. 1b (black). If the current quarantine measures continue, $\mathcal{R}(t) < 1$ is expected to be around the simulation day 71 (95% CI: 69–73), which indicates that the disease is controlled [12].

We also fitted the model with the available data for RJ. Considering the current social distancing policy, the projections for COVID-19 infection in RJ are showed in Fig. 2. According to Fig. 2a, the peak of active cases will occur on the simulation day 80 (95% CI: 76–85) with around 7.3 thousand (95% CI: 6.3–8.7) active cases, and D is expected to be around 7.0 thousand (95% CI: 5.8–8.7) in the considered time range. Fig. 2a also shows that C in the considered time range will be around 45.0 thousand (95% CI: 37.7–55.5). The $\mathcal{R}_0$ is 2.83 for RJ, and the distribution of $\mathcal{R}(t)$ along time



(a) Prediction of P, I, A, D, and C (right y-axis) in RJ. The vertical line indicates the curve peak, and the dashed lines display the corresponding 95% CI.

(b) Evolution of $\mathcal{R}(t)$ in RJ. Dashed lines indicate 95% CI around $\mathcal{R}(t) = 1$. The black curve refers to the simulation shown in Fig. 2a, and others to simulations with different distancing measures.

Figure 2: Dynamics of the COVID-19 in RJ modeled with the available data (Materials and Methods). The lines indicate the simulation with MAP estimates, with the shaded colors indicating the 95% CI. Simulation days 0 and 80 corresponds to March 10, 2020 and May 28, 2020, respectively.

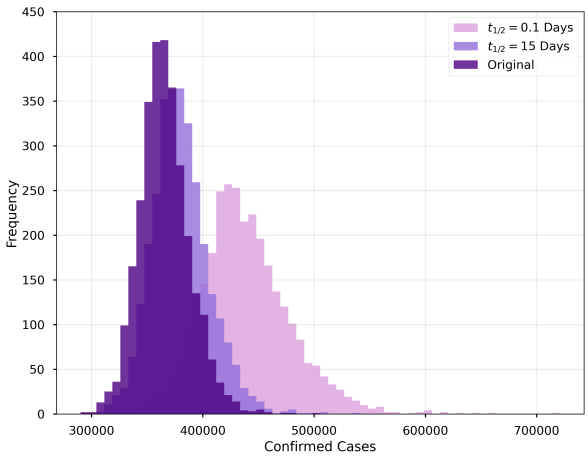
is displayed in Fig. 2b (black). Similarly as for the broader scale of the Brazilian scenario, if the quarantine measures in RJ continue, $\mathcal{R}(t) < 1$ is expected to be around the simulation day 72 (95% CI: 67–76).

Sensitivity analysis and influence of uncertainty in early COVID-19 epidemic

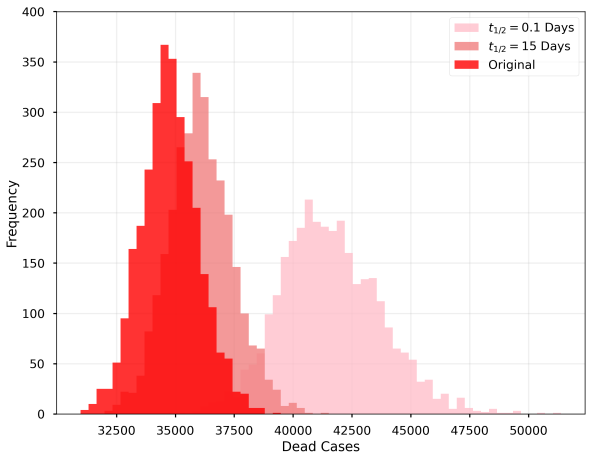
We address the sensitivity (Materials and Methods) for both C and D ($\text{QoI}_1(t)$), as well for $\mathcal{R}(t)$ ($\text{QoI}_2(t)$). Our first order sensitivity coefficient results suggest that uncertainties on ICs can drive significant changes in the early epidemic dynamics. However, the influence of such uncertainties diminishes as the epidemic evolves, indicating that other involved infectious mechanisms have more impact in the long term. Indeed, notice that parameter uncertainties have strong impact on both epidemic numbers ($\text{QoI}_1(t)$ and $\text{QoI}_2(t)$) along the dynamics. The importance of the recovery and the quarantine removal rates increase as the time evolves with respect to $\text{QoI}_1(t)$. Overall, the most influential parameter is the quarantine removal rate (ω), and its impact increases as the dynamics evolves. The second more influential parameter is the proportion of I with symptoms (ρ). We remark that the latter parameter is fixed due to the limited testing capacity in BR.

How would different social distancing measures affect the disease spread?

We have shown that the rate at which S, E, and I are removed due to quarantine measures significantly affect $\mathcal{R}(t)$, C, and D. Such average values under the actual social distancing policy are an idealization since actual values are dynamic and spatially heterogeneous. In order to study the qualitative effects of changes in social distancing policies, we propose to model ω as an exponential decay function, which can represent different social distancing relaxation scenarios (Materials and Methods). We consider three cases: (i) a sudden release from quarantine/social distancing, which is modeled as a decay with half-life as $t_{1/2} = 0.1$ days; (ii) a gradual release, with half-life as $t_{1/2} = 15$ days, and (iii) the current social distancing policy, with constant removal rate. Regarding such



(a) Posterior distributions of the maximum number of confirmed cases.



(b) Posterior distributions of the maximum number of dead cases.

Figure 3: Comparison of the posterior distributions of the maximum numbers of confirmed (a) and dead (b) cases for different social distancing scenarios in BR. The original scenario corresponds to keeping ω fixed at its MAP estimate. Analysis is done on the last simulation day (262).

idealized scenarios and considering that the beginning of the social distancing release policy is at simulation day 100 after the first time stamp considered in the calibration dataset, Fig. 3 shows the consequences of relaxing social distance measures in terms of C and D for BR. Additional results are displayed in SI-SAppendix A.6 and SI-SAppendix A.7 for both BR and RJ scenarios. Remarkably, a sudden release can induce an increase in C, extending the crisis duration due to a slower decrease in active cases. This effect can be noted in Figs. 1b and 2b. Social distancing release policies make the decrease of $\mathcal{R}(t)$ far slower, implying in a slow control of the disease. Thus, determining an appropriate moment to begin a social distancing release policy demands special care, since applying such policies in inappropriate times can maintain a crisis status for a longer-term. Moreover, due to the small changes in $\mathcal{R}(t)$ for a longer time, more cases of C and D will occur, worsening the health damage in the total population.

We also remark the importance of the quarantine release date in the following less favorable scenario. Considering that social distancing is released 20 days before the peak of active cases for BR (simulation day 61), disease spreading can yield a critical scenario even when a smooth and gradual release strategy (half-life as $t_{1/2} = 20$ days) is adopted, causing an increase in C and D, and a drastically longer epidemic period, as shown in Fig. 4. In this case, the peak of active cases will occur on the simulation day 95 (95% CI: 88–107) with around 85.7 thousand (95% CI: 73.4–105.0) active cases, and D is expected to be around 84.3 thousand (95% CI: 69.6–105.9) in the considered time range, with C around 0.8915 million (95% CI: 0.6834–1.246).

Discussion

As in other countries that adopted isolation measures, the spread of COVID-19 in BR had a slow start. Despite that, after about two months since the first confirmed case, on May 20, 2020 BR reaches second place among the countries that present more confirmed cases of the disease, only behind the United States of America [3]. The results shown in this work point to the necessity to quantify the effectiveness of social distancing interventions for the control of the disease in BR, specifically in its most affected localities. Besides, it is essential to take rapid responses to the emergence of new cases and their impact on the public health system. The Brazilian continental

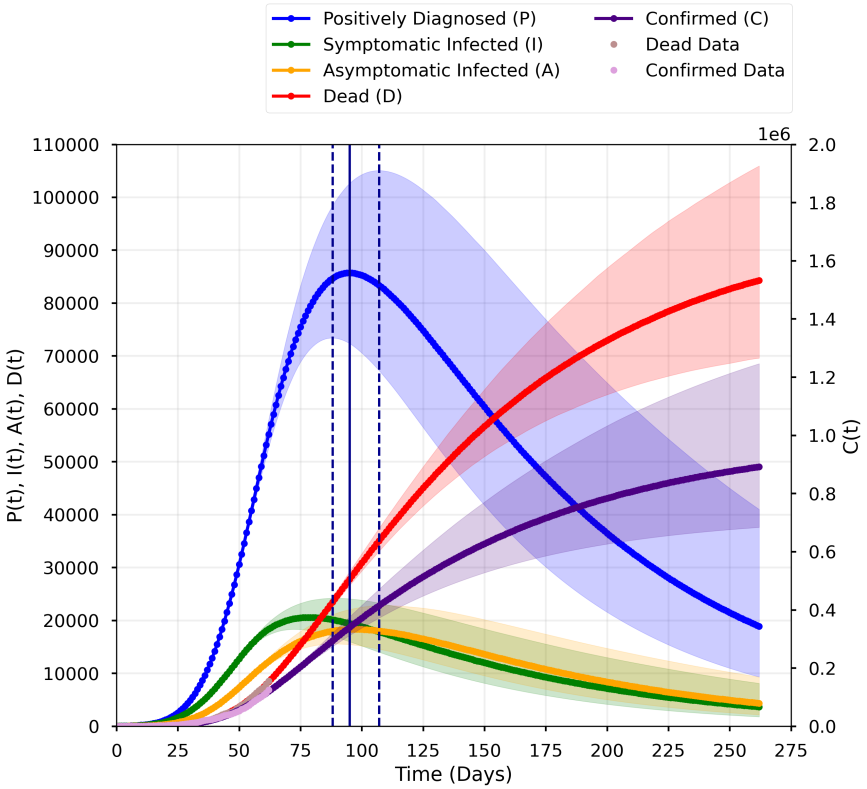


Figure 4: Prediction of P, I, A, D, and C (right y-axis) in BR, considering $t_{1/2} = 20$ days. Social distancing is released 20 days before the peak of active cases (simulation day 61).

dimension and great social and demographic differences are factors that hinder the elaboration of public policies for the disease control.

The present paper aims to contribute to the development of a model framework to investigate the expansion of COVID-19 in Brazilian locations and the impacts of different measures of social distancing. We apply the developed approach to model the COVID-19 dynamics in BR and RJ. This high populated Brazilian state was one of the first states to adopt mitigation actions, such as the suspension of classes, cancellation of events, and home isolation [13]. According to the current data released by the Brazilian Ministry of Health, RJ is one of the most affected states, both in the number of registered cases of COVID-19 and in the number of deaths. For this reason, the present study analyzes the spread in this state, as well as in the country as a whole, in order to assess the particular characteristics of the pandemic at those different spatial scales. Extensive research for the spread of COVID-19 in BR with multiple perspectives has been reported, e.g., [7, 13, 8]. The present work takes into account factors that are predominant in the Brazilian reality, such as the current limited testing capacity and the policy to test only severely ill hospitalized individuals. We hope that the present investigation brings some insights or guidelines for public health and policy-makers.

Due to data paucity, parameter identifiability is a major difficulty. To overcome this issue, only four model parameters were calibrated using a Bayesian approach. Other parameters, as well as model ICs, were set based on available information on COVID-19 (SI-Table A.3). Our simulation forecasts that the peak of active cases occurs on May 24, 2020 (95% CI: 21–26) and May 28, 2020 (95% CI: 24–2) for BR and RJ, respectively. This difference can be explained due to the discrepancy of the spatial scale, and the way the disease has spread along the different

Brazilian locations, considering, for instance, the different political measures taken and the social and demographic structure of each state. The social measures implemented in RJ at the beginning of the pandemic seemed to be able to flatten the epidemic curve and postpone the peak of active cases [13].

Aiming to evaluate the influence of uncertainties in hard-to-track populations such as undiagnosed infected individuals (I and A), as well those who carry the disease and are unable to transmit (E), we performed a sensitivity analysis to understand which model factors (parameters and ICs) play important roles at the various stages of the epidemic for both BR and RJ. The analysis confirms that a proper understanding of how the disease spreads can provide insights and aids to elaborate containment decisions in order to reduce $\mathcal{R}(t)$. In this sense, considering both BR and RJ, sensitivity analysis suggests that the most influential parameter for a long term perspective is the quarantine removal rate parameter (ω).

We have shown that the rate at which S, E, I and A individuals are removed due to quarantine measures significantly affects $\mathcal{R}(t)$, C and D. Such average values under the actual social distancing policy are an idealization since actual values are dynamic and spatially heterogeneous, as mentioned in the last paragraph. In order to study the qualitative effects of changes in social distancing policies, we propose to model ω as an exponential decay function, which can represent different social distancing relaxation scenarios. When more abrupt social distancing relaxation is implemented after the occurrence of the peak of active cases, it accompanies a longer extension of the duration of the disease, with approximately 19% increase in the projected numbers of C and D. If implemented before the peak, the consequences can be devastating, as indicated by our results. The hypothetical scenario built by considering a slow and gradual release implemented 20 days before the original peak indicates a delay of 14 days in the occurrence of the peak of active cases, with more than 890 thousand C and about 85 thousand D accumulated over less than nine months of the presence of the disease in BR. Our simulations highlight the importance of relaxing social distancing measures only under a very careful follow-up.

We note that the analysis performed in this paper should be viewed from a qualitative perspective. The conclusions for the considered hypothetical scenarios (with and without distancing relaxation) are based on model predictions and current employed policies. Model simplifications and the calibration procedure of model parameters can explain potential quantitative discrepancies between our predictions and data. Such simplifications include: (i) homogenization of age, social, and spatial structure, (ii) some model parameters are fixed, with values obtained from the literature, and (iii) model parameters are constant along time (except ω, that in some simulations varied in time). Moreover, data have limited information due to sub-notification, since only hospitalized cases are tested in BR. These simplifications are inherent of the modeling procedure, and should be viewed as part of the scientific process of the understanding of the natural phenomena. The present paper can be extended in forthcoming studies, for example, by considering models with spatial heterogeneity, parameter dependence in time and space, and also analysing data considering sub-notification. Data considering other BR states could also be analyzed employing the same procedure adopted in the present paper. Another possible extension could be related to the determination of limit thresholds in the number of P cases under different relaxation social distancing measures.

It is important to highlight the impacts of social distancing relaxation in order to control the epidemic. Adoption of this type of measure directly affects the evolution of $\mathcal{R}(t)$, as shown in our results. For both BR and RJ, it implies a much slower stagnation or decrease in $\mathcal{R}(t)$ from the time the measures are implemented. Since $\mathcal{R}(t)$ is in a controlled situation for both scenarios, i.e. $\mathcal{R}(t) < 1$, a direct consequence is that the disease would need more time to be eradicated. Our analyses suggest that policies based on short-term social distancing are not enough to control

the evolution of the pandemic. If social distancing policy measures are released before the optimal time, a second peak should be experienced [7]. Some authors argue that longer or even intermittent social distancing will be necessary to avoid recurrent outbreaks. Specifically, [14] examined a range of likely virus transmission scenarios until 2025 and assessed non-pharmaceutical interventions to mitigate the outbreak. They concluded that if the new coronavirus behaves in the same matter to similar viruses we can expect the disease to return in the coming years, depending on the level and duration of immunity, an aspect that remains to be clarified in the future. The discovery of a vaccine, or new treatments, coupled with the testing of the population could alleviate the need for severe social distancing measures to control the disease. Until then, the need to maintain, even if intermittently, social distancing measures must be carefully addressed.

Materials and Methods

The SEAIRPD-Q model

Here we consider a generalized SEIRD model that includes protective quarantine measures and whose development is based on the following assumptions: (i) the analysis time is small enough such that natural birth and death are disregarded; (ii) all positively diagnosed individuals are hospitalized. The reason is that only badly ill infected individuals are subject to tests in BR; (iii) only symptomatic infected and positively tested (diagnosed/hospitalized) individuals may die due to complications from the disease, (iv) the hospitalized individuals are under treatment and remain isolated, so that they are considered not infectious; (v) quarantine measures are restrictive so that quarantined individuals are isolated and are not likely to be infected; and (vi) the recovered/removed individuals acquire immunity.

The previous assumptions reflect our current knowledge of the COVID-19 and are the base of our SEAIRPD-Q model that has seven compartments, including the population of positively diagnosed (P) individuals who are under medical treatment, and the dead (D) individuals. The infected class is modeled as two separate compartments encompassing individuals with and without symptoms, denoted by I and A, respectively. The removed compartment includes the recovered individuals as well as those under social distancing measures. The mathematical description (differential equations) of the SEAIRPD-Q model and its schematic representation are shown in Fig. 5 while model parameters and related meanings are exhibited in Table 1 and SI-Table A.2. Information regarding the numerical differential equation solver employed can be seen in SI-SAppendix A.2.

Susceptible individuals become exposed to COVID-19 when in contact with infected people. The rate at which this happens may vary depending on whether the infected individual has symptoms or not. After an incubation period, exposed individuals become infected. The fraction of infected people who present symptoms can vary widely. Due to the reduced testing capacity in BR, asymptomatic individuals, as well as symptomatic infected individuals with mild symptoms, are not likely to be hospitalized and therefore will not be diagnosed. Those who present symptoms are either isolated or hospitalized. We assume that only symptomatic infected individuals with stronger symptoms are diagnosed quickly and require hospitalization. This assumption better reflects the Brazilian policy on only testing severely ill individuals. Following the WHO guideline on social distancing, most Brazilian localities are adopting social distancing policies at some extent. An interesting feature of our model is the assumption that susceptible, exposed, and infected individuals can be removed due to quarantine policy. Thus, the removed individuals' compartment includes individuals who have recovered from the disease as well as those put in quarantine. At a removal rate ω, quarantined individuals are removed and kept isolated. It is known that a quarantine period above 14 days is sufficient to recover exposed and infected individuals, who are idealized immune after such period. Asymptomatic and symptomatic individuals can recover

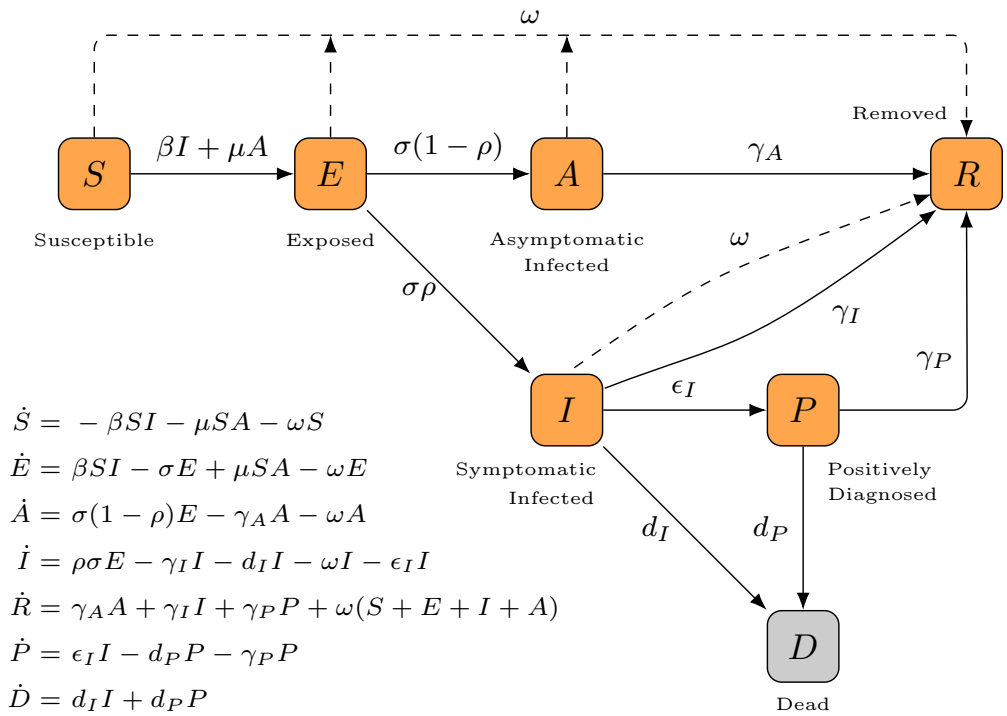


Figure 5: Mathematical formulation and schematic description of SEAIRPD-Q model.

Table 1: Model parameters: fixed values and MAP estimates of marginal distributions for BR and RJ (in appropriate units). The latter are obtained using a Markov Chain Monte Carlo (MCMC) like method, using the indicated prior distributions. A more detailed description of all variables, parameters, and ICs are shown in SI-SAppendix A.1.

Parameter	Value/MAP (BR)	Value/MAP (RJ)	Priori Distribution	Data Source
$\beta = \mu$	5.965935×10^{-9}	4.658646×10^{-8}	$\mathcal{U}(0, 1 \times 10^{-5})$	MCMC
σ	1/5	1/5	–	Table A.3
ρ	0.85	0.85	–	Table A.3
ε_I	1/3	1/3	–	Table A.3
$\gamma_A, \gamma_I, \gamma_P$	1/14	1/14	–	Table A.3
d_I	1.356770×10^{-2}	5.549912×10^{-4}	$\mathcal{U}(0, 0.1)$	MCMC
d_P	4.1681711×10^{-3}	1.315661×10^{-2}	$\mathcal{U}(0, 0.1)$	MCMC
ω	1.970400×10^{-2}	1.438746×10^{-2}	$\mathcal{U}(0, 1)$	MCMC
$\mathcal{R}_0$	4.093360	2.830033	–	SI-SAppendix A.3, (A.6)
σ_C	2040	148	$\mathcal{U}(1, 10000)$	MCMC
σ_D	81	22	$\mathcal{U}(1, 10000)$	MCMC

without medical treatment at rates γ_A and γ_I, respectively, and hospitalized individuals recover at a rate of γ_P.

Time dependent quarantine policy modeling

To represent and understand quarantine policy effects on COVID-19 epidemic, we propose a time-dependent removal rate, which we call ω_r. This function replaces the constant quarantine removal rate (ω), for each quarantine policy relaxation scenario. In this respect, we consider ω_r a smooth decreasing function $\omega_r = \omega e^{-\lambda(t-t_0)}$, with a decay constant $\lambda = \ln 2/t_{1/2}$, t_0 being the time at which the relaxation policy is implemented, and $t_{1/2}$ the half-life time for the quarantine release policy.

Bayesian calibration

Model calibration (SI-SAppendix A.4) is performed using a Bayesian approach. The code implementation is written in Python language (version 3.7), using PyMC3 as a Probabilistic Programming framework [15]. Due to the fact that solving ODE systems – such as SEAIRPD-Q proposed in the present work – can be computationally intensive, we choose as calibration method the Cascading Adaptive Transitional Metropolis in Parallel [16], a Transitional Markov chain Monte Carlo available in PyMC3. For the sake of reproducibility, code and scripts are publicly provided at [17].

As indicated in Table 1, we calibrated the following parameters: $\beta = \mu$, ω, d_I, and d_P. These parameters were fitted using cumulative data for C and D as observed quantities. The model likelihood function assumes a normal distribution considering these observed quantities as its mean values. The standard deviation for C (σ_C) and D (σ_D) are tuned as hyperparameters within the calibration procedure.

Data

Since Brazilian policy for testing in COVID-19 is still restricted to severe cases, model fitting aims at matching C and D, without accounting the recovered outflow from P. We gathered data for C and D from [18] (BR) and [19] (RJ). We assessed the epidemiological data of 63 days for BR, ranging from March 5, 2020 to May 6, 2020, and 53 for RJ, ranging from March 10, 2020 to May 1, 2020. These data ranges were chosen considering a minimal requirement of at least five individuals in the C compartment at the initial date.

Sensitivity analysis

In order to quantify how changes in the model parameters and some ICs affect certain quantities of interest (QoI), we apply the Elementary Effects method [20] present in the SALib library [21]. This method allows a global sensitivity analysis (SI-SAppendix A.5) by sampling initial points over all parametric space, computing for each of them how the QoI varies along generated trajectories (which are changes in the paramaters) and computing the associated sensitivity coefficients. The first order sensitivity coefficient of a certain parameter expresses how important it is to the considered QoI. The analyzed quantities are $\mathrm{QoI}_1(t) = \mathcal{R}(t)$ and the normalized sum of the squares of C and D, $\mathrm{QoI}_2(t) = \sqrt{C(t)^2 + D(t)^2}$. A 50% variation around the fixed and MAP values in Table 1 is considered to construct parameter ranges. Regarding the setup of the Elementary Effects procedure, we use a discretization level of $p = 4$ and a number of generated trajectories $m = 40$.

Acknowledgments

ACMR, LA, and JVOS would like to thank the support through CNPq projects 132591/2014-6, 301327/2020-3, and 166171/2018-2. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Finance Code 001.

Appendix A. Supporting Information

This Supporting Information (SI) text is organized as follows. Sections SI-SAppendix A.1, SI-SAppendix A.2, and SI-SAppendix A.3 give additional modeling details to complement the SEAIRPD-Q description reported in the main document. Sections SI-SAppendix A.4 and SI-SAppendix A.5 expand the Bayesian calibration and sensitivity analysis sections of the main text (Materials and Methods). Sections SI-SAppendix A.6 and SI-SAppendix A.7 expand the results section of the main document with additional figures for the BR and RJ scenarios, respectively. The last part of the document shows the calibration data used in this work.

Appendix A.1. Model Description

Mathematical models can help the investigation of different scenarios, devising alternative strategies to control virus spreading. The SARS coronavirus epidemic in 2002-2003 generated several publications using diversified modeling techniques (see [5] for a review). Some of these publications employed mathematical models to provide some understanding of the propagation pattern and ultimately to infer control measures for the disease [22, 23, 24]. Similar mathematical models were employed for other diseases, for example: (i) the 1918 influenza pandemic using data from cities of the USA [25, 26], England and Wales [27]; (ii) the 2009 influenza pandemic was modeled in local regions in Japan [28]; and (iii) the 2014-2016 Ebola epidemic was modeled for the West African countries with reported cases of the disease [29, 30]. Overall, the modeling efforts for representing virus spreading share common challenges that have been pointed out in the literature recently [31, 32], including those related to the emergence of "novel infectious agents" [33]. Issues like virus genetic diversity, within-host variability, social, demographic, and spatial heterogeneity, among others, can have a significant impact on model predictions and control programs. In any case, model granularity and ability to capture the desired phenomena are linked to available data for model calibration, which may hinder the development and practical use of more complex models.

Several mathematical models have already been published to better understand the disease spread and the effectiveness of different mitigation programs. Two variations of the Susceptible-Infectious-Recovered (SIR) model are used in [7] to describe the early evolution of COVID-19 in Brazil (BR), using data from February 25, 2020 to March 30, 2020. Short and long term forecasts are performed, which highlight the need of keeping social distancing policies for longer periods to actually flatten the epidemic curve, and the necessity of testing to have a better picture of the amplitude of the peak of symptomatic infected individuals. In [34], an eight compartment model with nine age stratification groups was developed to investigate the effect of social distancing policies, specifically those related to specific (older) age groups. Starting from a 60% lockdown scenario and many predefined parameterizations based on available references, their analysis indicates the ineffectiveness of older age group confinement policy as a way of controlling the confinement exit. A similar model approach is developed in [35] to investigate the transmission and isolation measures in São Paulo state. They found that both young and older people must be isolated to reduce hospitalization. The two latter studies explicitly included quarantined individuals, either due to confinement of susceptible individuals or because of a positive testing isolation measure. More recently, the work [36] develops a model that accounts for either reported and unreported cases. Besides analyzing the impact of different public health mitigation measures, the work suggests the use of a mixture approach for estimating model parameters. At the early onset of the disease, confirmed reported cases are assumed to follow an exponential law to ease getting initial conditions (ICs). Later on, a Bayesian approach is used for model calibration. Since the unreported infected cases are major issue to control outbreaks, the approach for identifying them seems promising and has already been successfully applied to the Hong Kong seasonal influenza epidemic in New York City in 1968–1969 for a SIR model [37].

Another class of compartmental models widely used to describe the spread of virus diseases worldwide is SEIR, that consider the total population (N) is split into sub-groups of Susceptible (S), Exposed (E), Infected (I), and Recovered or Removed (R) individuals, such that $S+E+I+R = N$. The compartments S and I contain all individuals likely to be infected and those already infected, respectively. E individuals are usually considered to have low viral load and are not likely to be infectious, while R individuals can or cannot be infected again depending on the virus disease. This class of dynamical model may be used to ultimately identify the processes that interplay in pandemic waves and has been largely used recently to assess the dynamics of COVID-19 in many countries [38, 39, 40, 41, 42, 43], including BR [44, 7, 34, 35, 36, 45, 46, 13].

Inspired in the generalized SEIR model proposed in [38], Jia et al. [47] employed the SEAIRPQ model with seven compartments for modeling the Chinese outbreak of COVID-19, including populations of positively diagnosed (P) individuals who are under medical treatment and of quarantined (Q) individuals. A schematic description of the SEAIRPQ is shown in Fig. A.6, with its mathematical formulation given by the following system of ordinary differential equations:

$$\begin{aligned}
\dot{S} &= -\beta S(I+\theta A) - pS + \lambda Q, \\
\dot{E} &= \beta S(I+\theta A) - \sigma E, \\
\dot{A} &= \sigma(1-\rho)E - \varepsilon_A A - \gamma_A A, \\
\dot{I} &= \sigma\rho E - \varepsilon_I I - \gamma_I I - d_I I \\
\dot{R} &= \gamma_A A + \gamma_I I + \gamma_P P, \\
\dot{P} &= \varepsilon_A A + \varepsilon_I I - \gamma_P P - d_P P, \\
\dot{Q} &= pS - \lambda Q.
\end{aligned} \tag{A.1}$$

Model parameters are non-negative constants and most of them are described in Table A.2. The three remaining are the fraction of infection rates between A and I, θ, the time required for an individual to be quarantined, $1/p$, and the quarantine period, $1/\lambda$. An interesting feature of the SEAIRPQ model is the explicit inclusion of a compartment for Q individuals. Longer periods of isolation yields a decrease in S population, ultimately decreasing the infected population. S individuals may be exposed to contact with A and I individuals. For the Chinese considered scenarios, it was assumed that A individuals are less likely to spread the disease, so that $\theta << 1$. The SEAIRPQ model assumes that both A and I individuals are eventually positively diagnosed at rates ε_A and ε_I, respectively. They can also recover without medical treatment at rates γ_A and γ_I, respectively.

Although SEAIRPQ model was able to describe quite well the outbreak in China, the Brazilian reality may prevent its use in Brazilian localities. BR testing capacity is still very limited, one of the lowest in the world. In April 2020, BR tested around 296 per million people. According to [3], this number increased to 4,104 in May 2020, although it can be significantly lower in some Brazilian localities due to huge demographic and social heterogeneities. Thus, available data on testing is very scarce: the present Brazilian policy is to test only people with severe symptoms. Such evidence has driven the development of the model proposed in this paper.

Here we consider a modification of model (A.1) taking into account protective quarantine measures as an implicit variable. The proposed model, called SEIARPD-Q, has also seven compartments since it includes the compartment of dead (D) individuals to explicitly keep track of them. The notation used indicates that Q individuals are not separated in a specific compartment. Instead, quarantine mechanisms are implicitly considered, with individuals in isolation included in R compartment. Likewise SEIARPQ model, the infected class is modeled as two separate compartments encompassing I and A individuals. R compartment includes individuals who have recovered from the disease as well as those who have been removed to this compartment as a result of isolation measures.

For completeness, we repeat here the mathematical formulation of the SEAIRPD-Q model, which is given by the following set of differential equations:

$$
\begin{aligned}
\dot{S} &= -\beta SI - \mu SA - \omega S, \\
\dot{E} &= \beta SI - \sigma E + \mu SA - \omega E, \\
\dot{A} &= \sigma(1-\rho)E - \gamma_A A - \omega A, \\
\dot{I} &= \rho\sigma E - \gamma_I I - d_I I - \omega I - \epsilon_I I, \\
\dot{R} &= \gamma_A A + \gamma_I I + \gamma_P P + \omega(S + E + I + A), \\
\dot{P} &= \epsilon_I I - d_P P - \gamma_P P, \\
\dot{D} &= d_I I + d_P P,
\end{aligned}
\tag{A.2}
$$

The model is schematically shown in Fig. A.7. All model parameters meanings are described in Table A.2. Data sources are informed in Table A.3 and Table 1 (Materials and Methods). Information regarding the numerical differential equation solver employed can be seen in SI-SAppendix A.2.

Looking at the mathematical description of model (A.2), it is important to note that the transformation of S into E individuals occurs through interaction with individuals from the infected groups (I and A) according to β and μ rates, respectively. The remaining mechanisms are linear, depending only on the corresponding transfer rates. As already mentioned, an important mechanism introduced in the model describes the application of social distancing measures: individuals from compartments S, E, A, and I are removed to compartment R at a rate ω. In this way, it is possible to exclude R individuals from the infection dynamics. A similar approach was adopted in [11]. We emphasize that we have chosen an implicit quarantine/social distancing approach, that is, we do not model a specific compartment of Q individuals, unlike other studies [47, 38]. The advantage of this implicit representation is to obtain a simpler model, with less computational cost and fewer parameters.

The previous assumptions reflect our current knowledge of the COVID-19 in BR. We remark that the adaptive immunity to SARS-CoV-2 of people who recovered from the disease has not been established yet as well as how long the immunity could eventually last. Actually, a follow-up study on 262 recovered patients in China revealed that 15% were tested positively for SARS-Cov-2. They were all younger than 14-years old and had mild and moderate conditions at the time of the primary infection. During the follow-up, they did not show clinical symptoms and it is speculated if either more sensitive test is required or indicates a sign of relapse rather than reinfection. Another study performed in previously infected rhesus macaques found that reinfection of SARS-Cov-2 could not occur [48]. Overall, the current understanding, based on what is known on coronaviruses in general, is that antibodies produced by infected patients may give them immunity to the specific virus for short term or even years. For example, previously infected SARS patients can become susceptible to reinfection only after around three years [49].

Appendix A.2. Ordinary Differential Equations Solving Method

The mathematical formulation of the SEAIRPD-Q model described by the system of seven ordinary differential equations (A.2) is solved using the LSODA method [50] from SciPy [51]. This package provides an interface for such routine from ODEPACK [52].

Appendix A.3. Evaluation of the Reproduction Number

Following the ideas introduced in [53], we apply the Next Generation Matrix (NGM) method to obtain $\mathcal{R}_0$ for the SEAIRPD-Q model. Firstly, we use a reduction process to arrive at a linearized infection model, under the disease-free steady state. The identification of the infected variables (E, A, and I) allows defining the vector $\mathbf{x}^T = [E, I, A]$, where the superscript T stands for the

transposition of a matrix. Thus, we consider that S, R, P, and D individuals are not able to transmit the disease. Denoting by T and Σ the transmission and transition matrices, respectively, the three-dimensional linearized infection sub-system is given by:

$$\frac{d\mathbf{x}}{dt} = (T + \Sigma)\mathbf{x}, \tag{A.3}$$

where

$$T = \begin{bmatrix} 0 & \beta S & \mu S \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{A.4}$$

$$-\Sigma = \begin{bmatrix} \sigma + \omega & 0 & 0 \\ -\sigma\rho & \gamma_I + d_I + \omega + \epsilon_I & 0 \\ -\sigma(1-\rho) & 0 & \gamma_A + \omega \end{bmatrix}. \tag{A.5}$$

The reproduction number $\mathcal{R}_0$ is the dominant eigenvalue of the next generation matrix $K = -T\Sigma^{-1}$ [53], which yields:

$$\mathcal{R}_0 = \left(\frac{\sigma(1-\rho)\mu}{(\sigma+\omega)(\gamma_A+\omega)} + \frac{\beta\sigma\rho}{(\sigma+\omega)(\gamma_I+d_I+\omega+\epsilon_I)} \right) S(0). \tag{A.6}$$

The corresponding effective reproduction number is:

$$\mathcal{R}(t) = \left(\frac{\sigma(1-\rho)\mu}{(\sigma+\omega)(\gamma_A+\omega)} + \frac{\beta\sigma\rho}{(\sigma+\omega)(\gamma_I+d_I+\omega+\epsilon_I)} \right) S(t). \tag{A.7}$$

Appendix A.4. Bayesian Calibration

In this work we adjust our model parameters to make results consistent with available observations (i.e., C and D for both BR and RJ scenarios). With Bayesian calibration, we are capable of determining the most likely uncertainties for input parameters as well as their MAP values described by the *a posteriori* probability distribution π_{post}. Considering that our aim is to adjust a set of parameters $\boldsymbol{\theta}$ given $\boldsymbol{y}$ observations, Bayes' theorem states:

$$\pi_{\text{post}}(\boldsymbol{\theta}|\boldsymbol{y}) = \frac{\pi_{\text{like}}(\boldsymbol{y}|\boldsymbol{\theta})\pi_{\text{prior}}(\boldsymbol{\theta})}{\pi_{\text{evid}}(\boldsymbol{y})}, \tag{A.8}$$

where π_{prior} is the *a priori* probability distribution that represents the initial (*a priori*) knowledge over $\boldsymbol{\theta}$, π_{like} is the likelihood function, and π_{evid} is the evidence (information) encompassed in the data (see [54] for a more detailed description). We assume that all data follows a Gaussian distribution centered on simulated of observable quantities values with variances σ_C^2 and σ_D^2, which yields the following form for the likelihood function:

$$\pi_{\text{like}}(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{j\in\{C,D\}} \frac{1}{\sigma_j\sqrt{2\pi}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} \left(\frac{y^{(j)}(t_i) - y^{(j)}_{\text{model}}(t_i)}{\sigma_j} \right)^2 \right) \tag{A.9}$$

The standard deviations for the cumulative confirmed (σ_C) and dead cases (σ_D) are tuned as hyperparameters within the calibration process. Regarding π_{prior}, a list of prior distributions and MAP estimates of marginal distributions can be seen in Table 1 (Materials and Methods). The frequency histograms for the calibrated parameters of both BR and RJ scenarios are shown in Figs. A.8 and A.12, respectively. The posterior distributions are obtained by a tempered adaptive

Bayesian updating procedure proposed in [16], the Cascading Adaptive Metropolis in Parallel, used here through PyMC3 [15] implementation in Python language (version 3.7).

Despite the fact that we have mentioned which observed data are used as observations for the calibration procedure, a proper definition of such quantities is absent. The estimate of C is calculated from the model as:

$$C(t) = \int_{t_0}^{t} \varepsilon_I I(t) dt, \tag{A.10}$$

which is the cumulative value of inflow individuals in the P compartment, since only diagnosed individuals are recorded on Brazilian confirmed case databases due to the lack of large-scale testing in the overall population. Using Eq. A.10, we can estimate the amount C at t days after day t_0. Analogously, in order to compute the D from SEAIRPD-Q, we calculate:

$$D(t) = \int_{t_0}^{t} \left[d_I I(t) + d_P P(t) \right] dt, \tag{A.11}$$

which is the integration of $\dot{D}$ (see system (A.2) for reference) over the time range of interest. C and D time-stamps are gathered from data sources specified in the Materials and Methods section. Such quantities are used as observed data in the calibration procedure.

Appendix A.5. Sensitivity Analysis

Due to paucity data and possible model identifiability issues, not every parameter and unknown initial conditions (ICs) are estimated using the Bayesian calibration procedure. Based on previous studies, we assume values for every parameter (with the exception of β, μ, ω, d_I, and d_P), and IC (with the exception of $S(0)$, $P(0)$, $D(0)$, and $R(0)$). Relevant information on model settings is provided in Tables 1 (Materials and Methods), A.3 and A.2. In order to quantify how these choices affect the outcome of the dynamics, we perform a sensitivity analysis. Specifically, we are interested in the sensibility of the quantities:

$$\mathrm{QoI}_1(t) = \mathcal{R}(t), \tag{A.12}$$

$$\mathrm{QoI}_2(t) = \sqrt{C(t)^2 + D(t)^2}. \tag{A.13}$$

As previously stated in the main text, we employ the Elementary Effects method [20] to carry out this analysis. In contrast to local sensitivity methods, which evaluate how small changes around a singular point affect a QoI [55], the Elementary Effects method is a global sensitivity method. Considering the $\boldsymbol{\theta} \in \mathbb{R}^k$ as the set of analysed parameters, we construct a sample of m initial points $\boldsymbol{\theta}^{(1)}, \cdots, \boldsymbol{\theta}^{(m)}$ from a k-dimensional p-level grid. Afterwards, a metric defined as an elementary effect EE_i, $i = 1, \ldots, k$, is computed for every parameter:

$$EE_i(\boldsymbol{\theta}) = \frac{\mathrm{QoI}(\theta_1, \ldots, \theta_{i-1}, \theta_i + \Delta, \theta_{i+1}, \ldots, \theta_k) - \mathrm{QoI}(\boldsymbol{\theta})}{\Delta}. \tag{A.14}$$

where $\Delta \in \{\frac{1}{p-1}, ..., 1 - \frac{1}{p-1}\}$ is the magnitude of change in the ith parameter in the k-dimensional p-level grid. Values of p as an even number and $\Delta = \frac{p}{2(p-1)}$ are usual suggested choices to ensure an equal sampling probability in the parametric space, and are the default option in the SALib [21] library used. After computing the elementary effects, the process ends by calculating the global sensitivity coefficients for each parameter. Our work dedicates the analysis only to the first order sensitivity coefficient:

$$s_i = \frac{1}{m} \sum_{j=1}^{m} EE_i(\boldsymbol{\theta}^{(j)}). \tag{A.15}$$

Parameters with high scores are the most influential to the QoI. In turn, the first sensitivity index values allow ranking the parameters with respect to their order of importance. Sensitivity analysis is performed for both scenarios (BR and RJ) considering all parameters and ICs $E(0)$, $A(0)$, and $I(0)$. The settings $p = 4$ and $m = 40$ are used for all experiments. A 50% variation around the fixed and MAP values in Tables 1 (Materials and Methods) and A.3 is considered to construct parameter ranges. The QoIs are observed in days 2, 25, 50, 75, 100, 125, 150, 175, 200, 225, and 250 in order to evaluate the scores s_i along all simulation period. For the sake of presentation, the scores are normalized (sum to 1).

The sensitivity analysis results for BR are depicted in Fig. A.10. The sensitivity indexes for the cumulative C and D ($\text{QoI}_2(t)$) are heavily influenced by the initial conditions in the early stages of the disease propagation, although their influence decreases as time evolves. In contrast, they are not influential for $\mathcal{R}(t)$ ($\text{QoI}_1(t)$) along all the simulation time. The influence of the parameter related to social distancing (ω) on both QoIs is notable. More importantly, its influence increases with the evolution of time, which emphasizes the need for care when establishing social distancing relaxation measures. On the other hand, the proportion of I individuals is also one of the most influential parameters on both QoIs, although it is more relevant considering C and D. This points out the need of having a more widespread testing policy. Regarding the RJ scenario, Fig. A.14 reveals very similar scores. One noteworthy distinction is the rank of influence of $E(0)$ and $I(0)$ associated with $\text{QoI}_2(t)$. The former is more influential in the BR scenario (Fig. A.10b), while the latter has a higher score for RJ (Fig. A.14b).

Appendix A.6. BR: Additional Results

Fig. A.9a shows the prediction of P, I, A, D, and C in BR. A total of 75.7 thousand (95% CI: 67.7–85.8) active cases are expected on the simulation day 81 (95% CI: 78–83), with D and C expected to be around 34.7 thousand (95% CI: 32.4–37.3) and 364.2 thousand (95% CI: 325.7–414.8), respectively. The posterior distribution of the peak position is displayed in Fig. A.9b, with the vertical dashed lines corresponding to those displayed in Fig. A.9a. Fig. A.9c depicts the temporal changes of $\mathcal{R}(t)$ and displays two vertical lines identifying the credible interval (95% CI: 69–73) of the time above which $\mathcal{R}(t) \leq 1$. The same lines are depicted in Fig. A.9d that shows the uncertainty associated with that time value.

Fig. A.11 provides the model forecasts for the considered hypothetical scenarios of social distancing relaxation. Fig. A.11a shows what to expect in case of a sudden release from quarantine after the simulation day 100 ($t_{1/2} = 0.1$ days as half-life decay). In this case, C reaches 434.8 thousand (95% CI: 372.1–526.8) and D 41.4 thousand (95% CI: 37.8–46.0) at the end of the simulation (day 262), which corresponds to an increase of approximately 19% in both values when compared to the original scenario with ω fixed at its MAP estimate. For a gradual release after the simulation day 100 ($t_{1/2} = 15$ days as half-life decay), the numbers of C and D at the end of simulation are 378.0 thousand (95% CI: 334.7–436.0) and 36.0 thousand (95% CI: 33.6–38.9), respectively, as shown in Fig. A.11b. Figs. A.11c and A.11d compare the variability of C and D at the end of simulation in the form of box plots for all scenarios considered. Outliers appear as individual points and the samples medians are depicted in red.

Appendix A.7. RJ: Additional Results

Fig. A.13a shows the prediction of P, I, A, D, and C in RJ. A total of 7.3 thousand (95% CI: 6.3–8.7) active cases are expected on the simulation day 80 (95% CI: 76–85), with D and C expected to be around 7.0 thousand (95% CI: 5.8–8.7) and 45.0 thousand (95% CI: 37.7–55.5), respectively. The posterior distribution of the peak position is displayed in Fig. A.13b, with the vertical dashed lines corresponding to those displayed in Fig. A.13a. Fig. A.13c depicts the temporal changes of

$\mathcal{R}(t)$ and displays two vertical lines identifying the credible interval (95% CI: 67–76) of the time above which $\mathcal{R}(t) \leq 1$. The same lines are depicted in Fig. A.13d that shows the uncertainty associated with that time value.

Fig. A.15 provides the model forecasts for the considered hypothetical scenarios of social distancing relaxation. Fig. A.15a shows what to expect in case of a sudden release from quarantine after the simulation day 100 ($t_{1/2} = 0.1$ days as half-life decay). In this case, C reaches 57.1 thousand (95% CI: 44.4–80.1) and D 8.8 thousand (95% CI: 6.8–12.3) at the end of the simulation (day 252), which corresponds to an increase of approximately 27% in both values when compared to the original scenario with ω fixed at its MAP estimate. For a gradual release after the simulation day 100 ($t_{1/2} = 15$ days as half-life decay), the numbers of C and D at the end of simulation are 47.8 thousand (95% CI: 39.3–61.1) and 7.4 thousand (95% CI: 6.0–9.5), respectively, as shown in Fig. A.15b. Figs. A.15c and A.15d compare the variability of C and D at the end of simulation in the form of box plots for all scenarios considered. Outliers appear as individual points and the samples medians are depicted in red.

Table A.2: Model parameters.

Parameter	Meaning
β	Rate of transmission per contact with symptomatic infected individual
μ	Rate of transmission per contact with asymptomatic infected individual
σ	Transition rate from exposed to infected individuals
ρ	Proportion of infected individuals who have symptoms
ε_I	Diagnosis rate of hospitalized individuals
γ_A	Recovery rate of asymptomatic infected individuals
γ_I	Recovery rate of symptomatic infected individuals
γ_P	Recovery rate of hospitalized individuals
d_I	Death rate of hospitalized individuals
d_P	Death rate of symptomatic infected individuals
ω	The rate at which susceptible, exposed, and infected individuals are removed due to quarantine measures

Table A.3: Fixed values for estimated parameters and ICs.

Parameter	Average Value	Hypothesis
ρ	0.85	The estimated asymptomatic proportion was 15% [56]
σ	1/5 day^{-1}	Estimated incubation (or latent) period based on [57]
ε_I	1/3 day^{-1}	We adopt that only severely ill individuals are diagnosed in BR, which takes about 3 days
γ_A	1/14 day^{-1}	It takes around 14 days for recovering from COVID-19 [58]
γ_I	1/14 day^{-1}	
γ_P	1/14 day^{-1}	
Class	**Initial Population (Ind.)**	**Hypothesis/Source**
N(0)	210147125 (BR), 17264943 (RJ)	IBGE 2019
P(0)	7 (BR), 8 (RJ)	Data at day 2020-03-05 (BR), 2020-03-10 (RJ)
E(0)	$10 \cdot$ P(0)	The number of exposed individuals is about 10 times the positively confirmed ones
I(0)	$5 \cdot$ P(0)	The number of symptomatic infected individuals is about 5 times the positively confirmed ones
A(0)	P(0)	The number of asymptomatic infected individuals is about the same of the positively confirmed ones
D(0)	0	Data at day 2020-03-05 (BR), 2020-03-10 (RJ)
R(0)	0	Data at day 2020-03-05 (BR), 2020-03-10 (RJ)
S(0)	210147006 (BR), 17264837 (RJ)	S(0) = N(0) - {E(0) + A(0) + I(0) + P(0) + R(0) + D(0)}

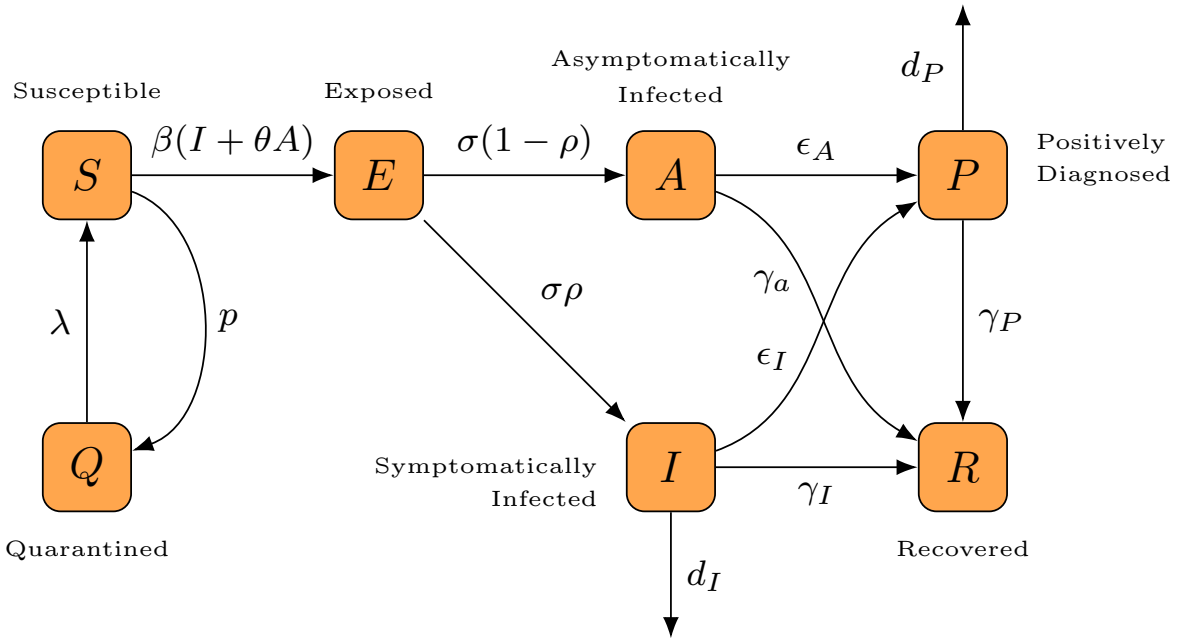


Figure A.6: A schematic description of the SEAIRPQ model, modified from [47].

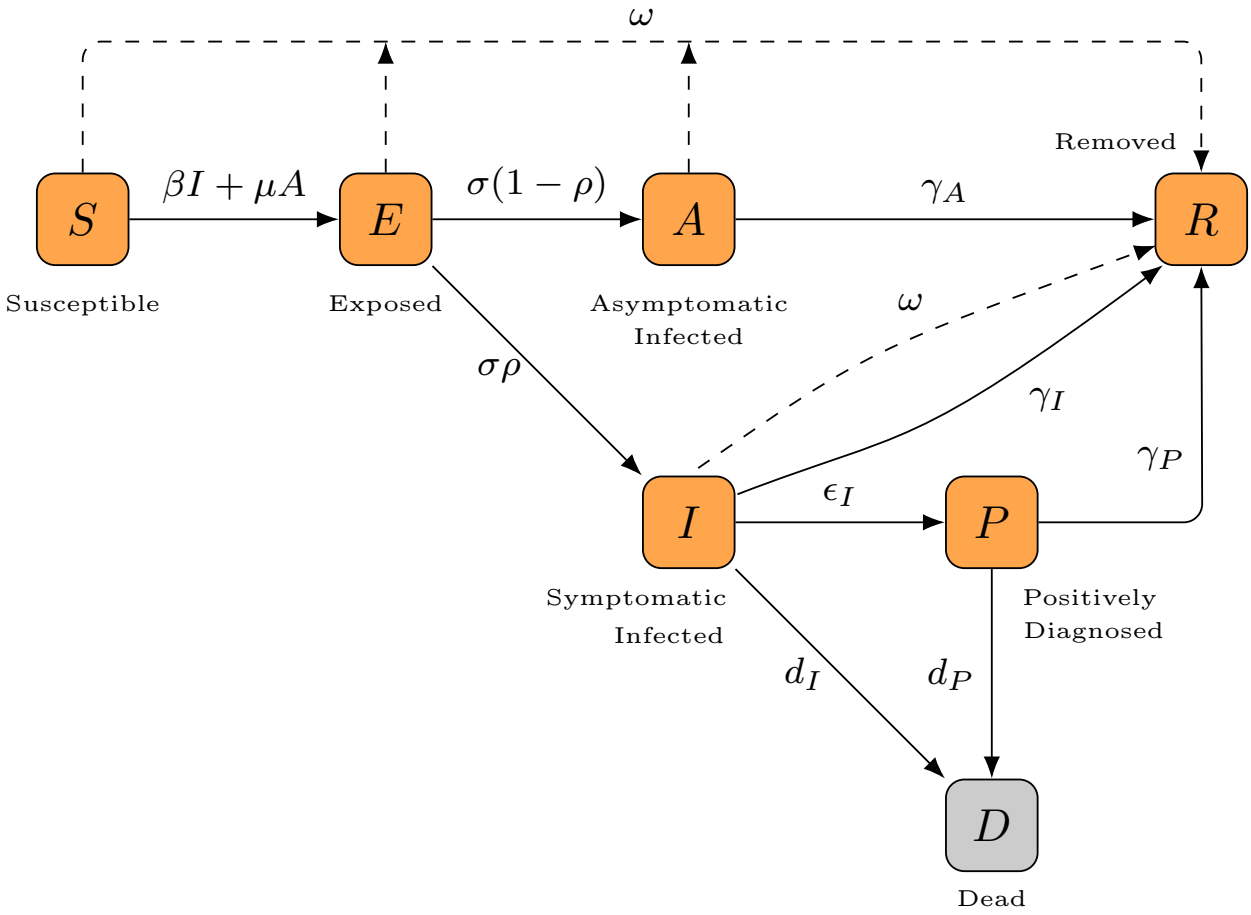
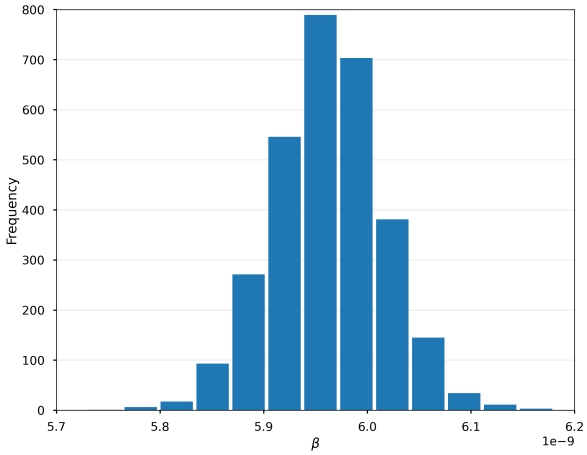
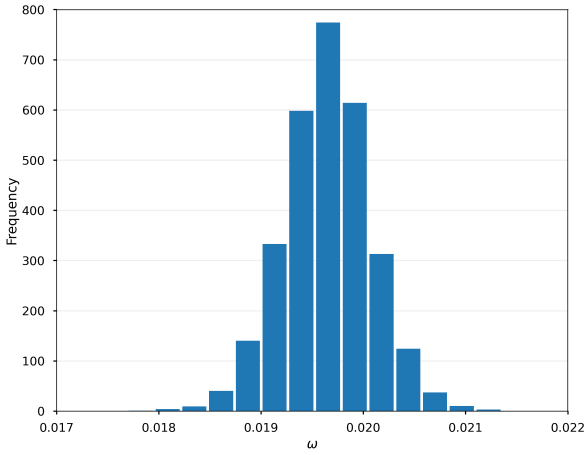


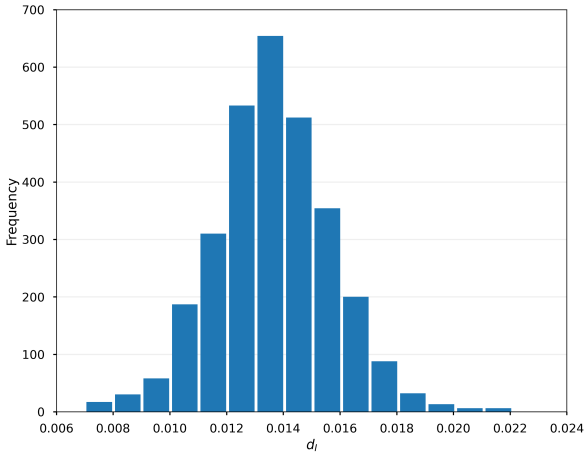
Figure A.7: A schematic description of the SEIARPD-Q model.



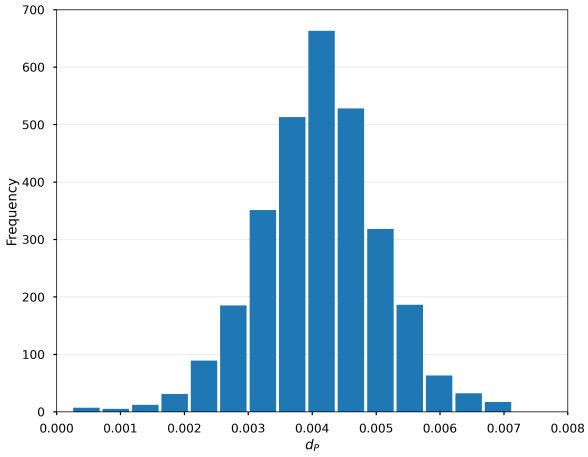
(a) Calibrated parameter: β.



(b) Calibrated parameter: ω.

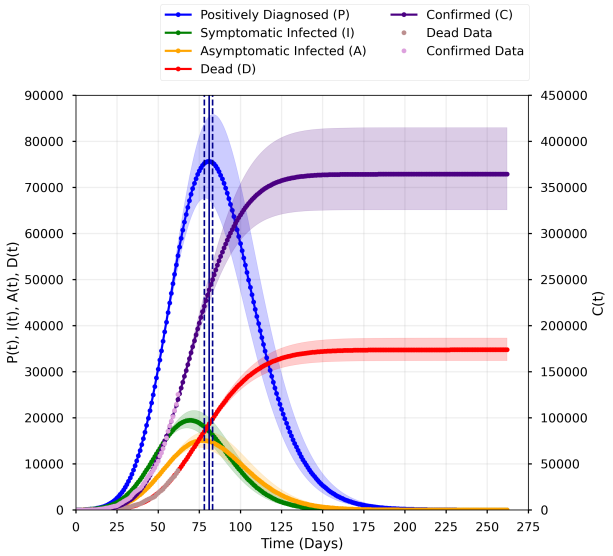


(c) Calibrated parameter: d_I.

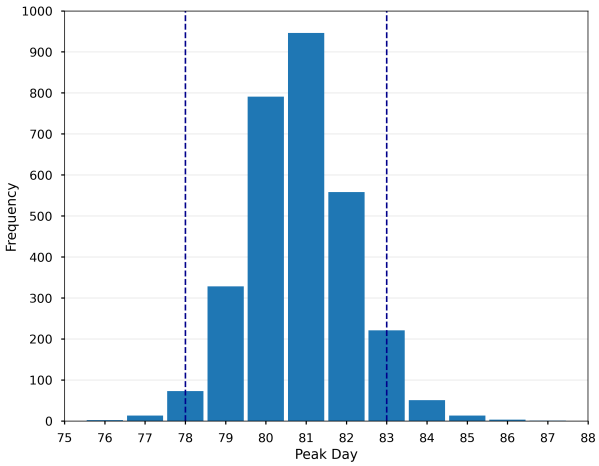


(d) Calibrated parameter: d_P.

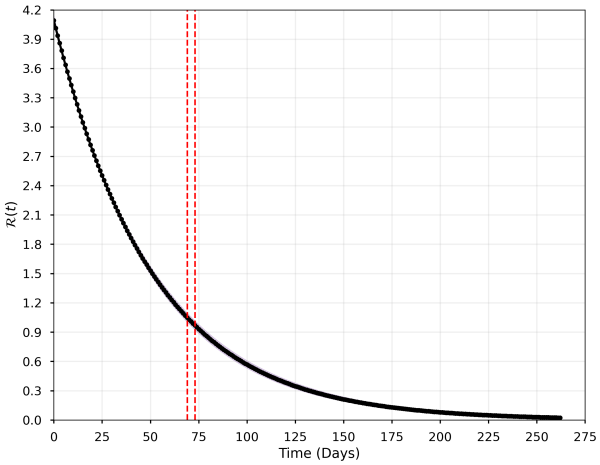
Figure A.8: Frequency histograms for the calibrated parameters (BR).



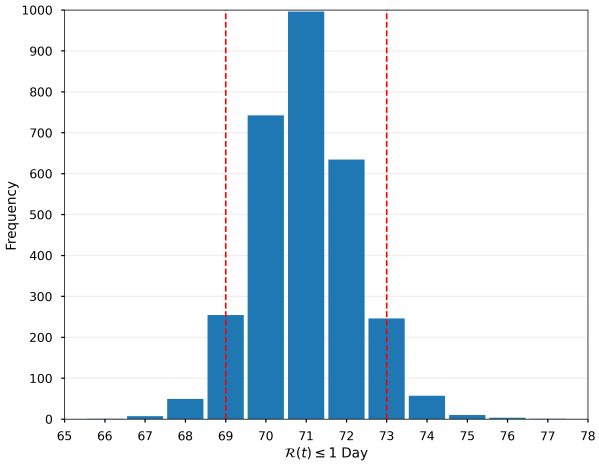
(a) Prediction of P, I, A, D, and C (right y-axis) in BR. The vertical line indicates the curve peak, and the dashed lines display the corresponding 95% CI.



(b) Frequency histogram for the peak day of P.

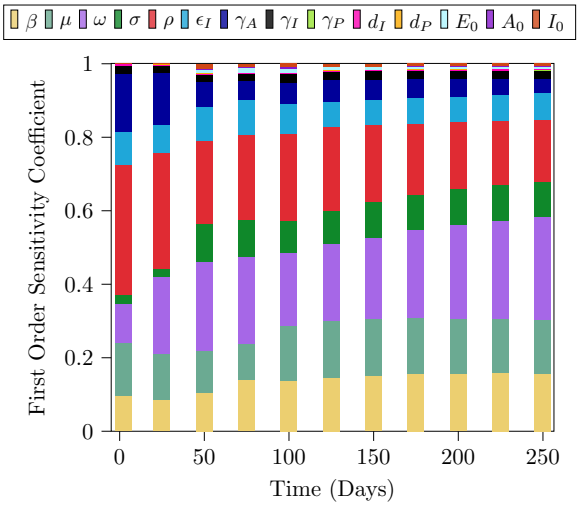


(c) $\mathcal{R}(t)$ of the COVID-19 in BR.

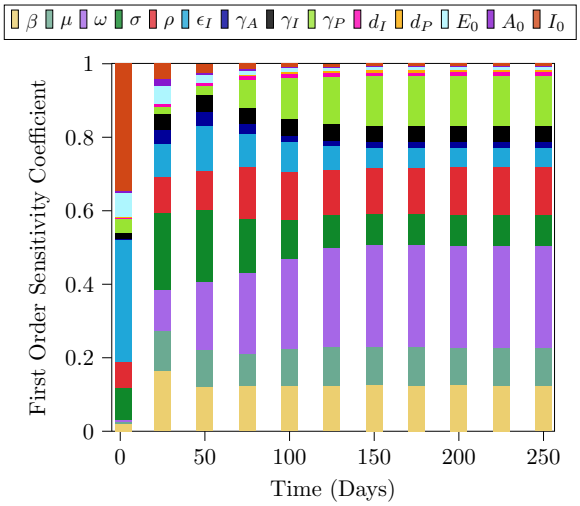


(d) Frequency histogram for the day at which $\mathcal{R}(t) \leq 1$.

Figure A.9: Dynamics of the COVID-19 in BR modeled with the available data (Materials and Methods).

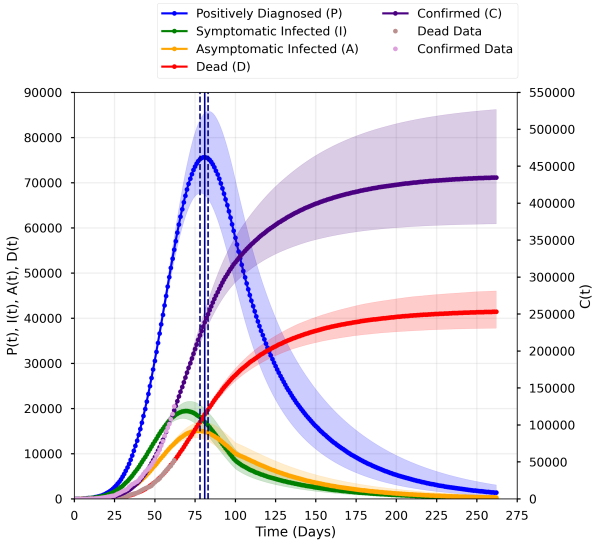


(a) Sensitivity analysis for $\mathcal{R}(t)$, $\text{QoI}_1(t)$.

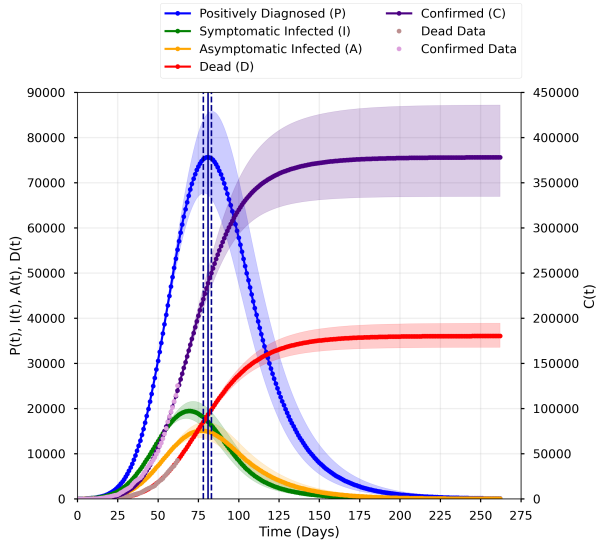


(b) Sensitivity analysis for C and D, $\text{QoI}_2(t)$.

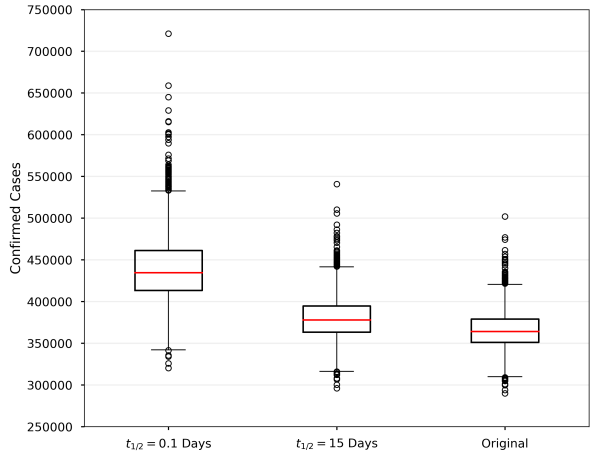
Figure A.10: Temporal changes of the first order sensitivity index of model factors for the BR scenario (Materials and Methods).



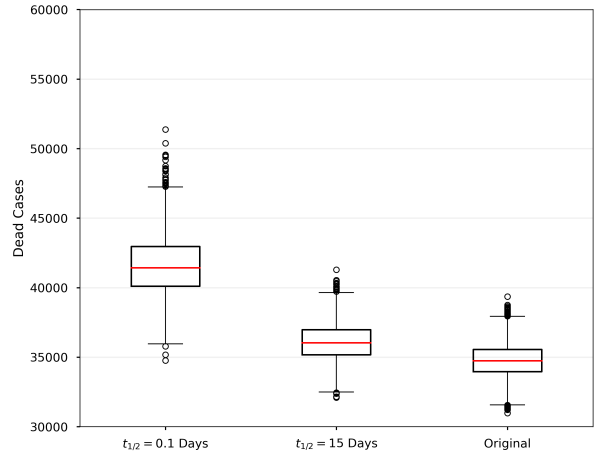
(a) Prediction of P, I, A, D, and C (right y-axis) in BR considering a sudden release from quarantine ($t_0 = 100$ days and $t_{1/2} = 0.1$ days). The vertical line indicates the curve peak, and the dashed lines display the corresponding 95% CI.



(b) Prediction of P, I, A, D, and C (right y-axis) in BR considering a gradual release from quarantine ($t_0 = 100$ days and $t_{1/2} = 15$ days). The vertical line indicates the curve peak, and the dashed lines display the corresponding 95% CI.

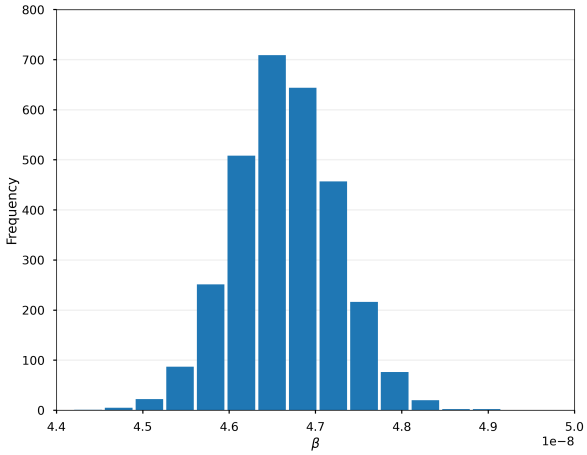


(c) Box plot for C in BR at the end of the simulation.

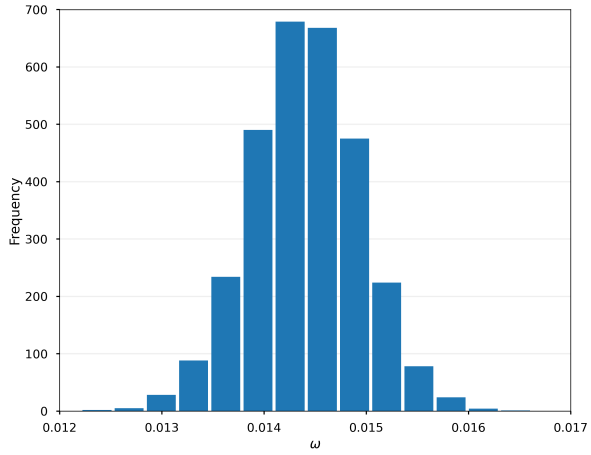


(d) Box plot for D in BR at the end of the simulation.

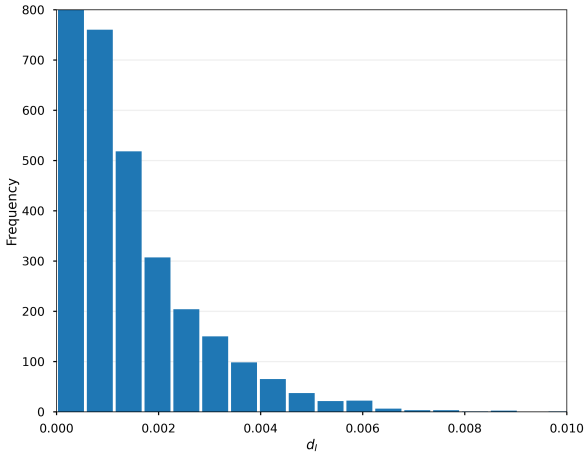
Figure A.11: Model forecasts in BR considering sudden and gradual releases from quarantine ($t_0 = 100$ days and $t_{1/2} = 0.1$ and 15 days). The original (baseline) scenario of the main text is also presented to ease comparison.



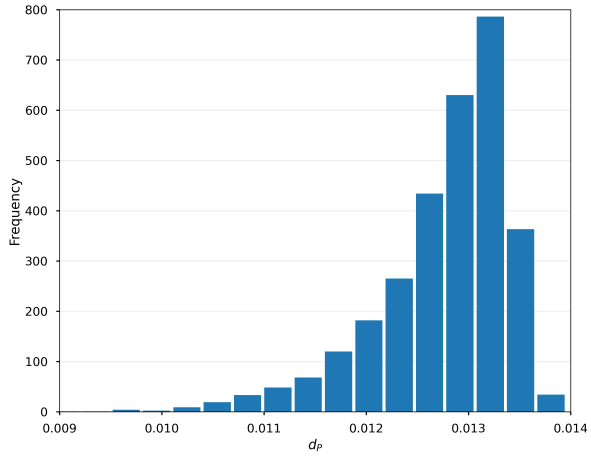
(a) Calibrated parameter: β.



(b) Calibrated parameter: ω.

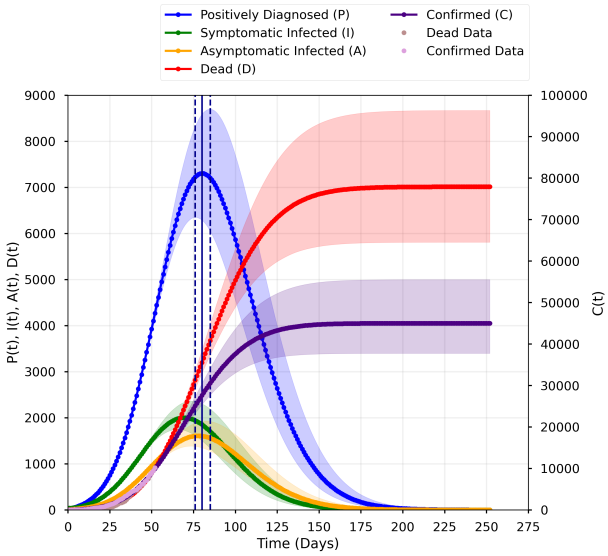


(c) Calibrated parameter: d_I.

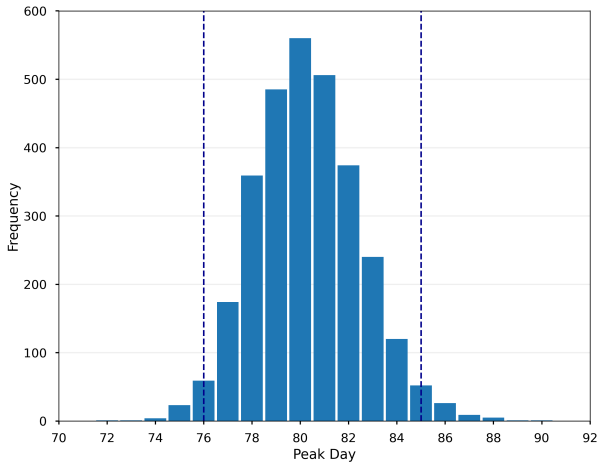


(d) Calibrated parameter: d_P.

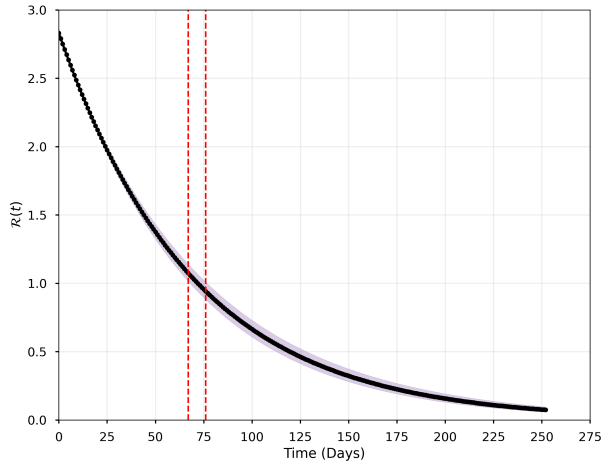
Figure A.12: Frequency histograms for the calibrated parameters (RJ).



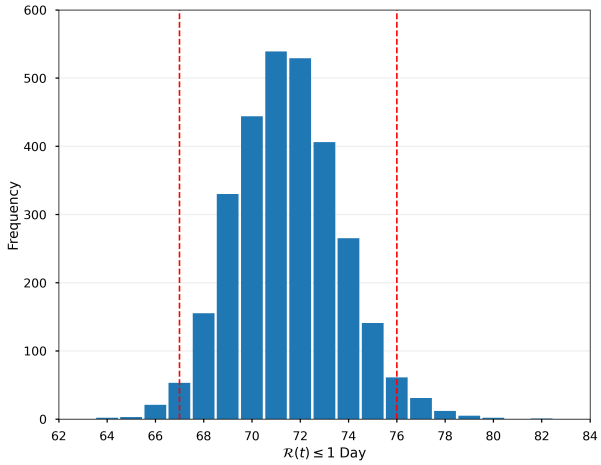
(a) Prediction of P, I, A, D, and C (right y-axis) in RJ. The vertical line indicates the curve peak, and the dashed lines display the corresponding 95% CI.



(b) Frequency histogram for the peak day of P.

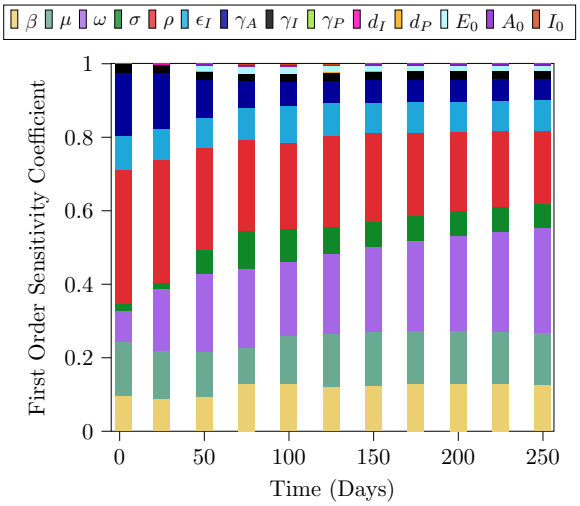


(c) $\mathcal{R}(t)$ of the COVID-19 in RJ.

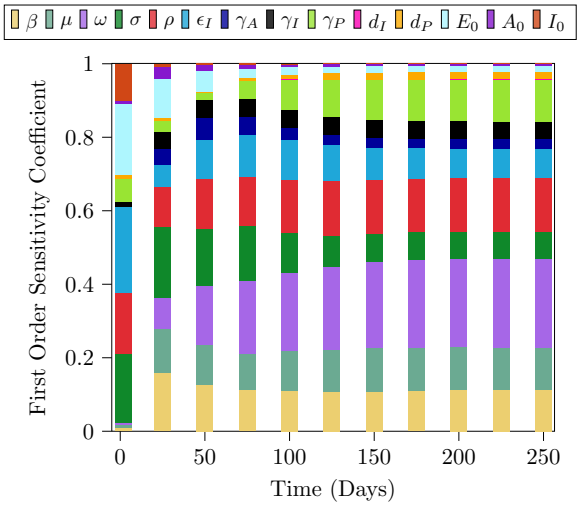


(d) Frequency histogram for the day at which $\mathcal{R}(t) \leq 1$.

Figure A.13: Dynamics of the COVID-19 in RJ modeled with the available data (Materials and Methods).

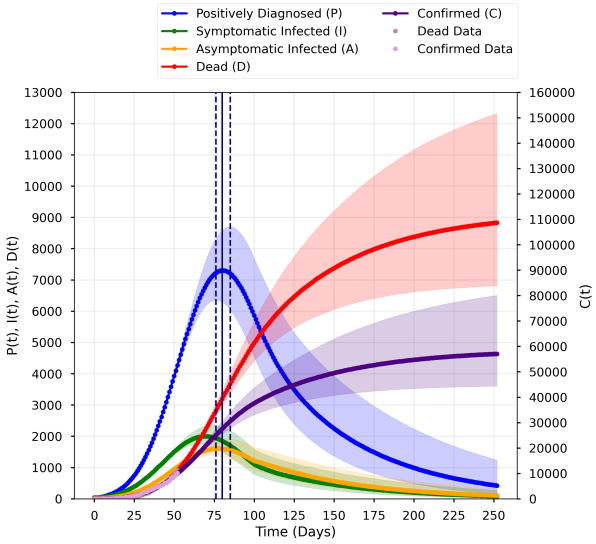


(a) Sensitivity analysis for $\mathcal{R}(t)$, $\text{QoI}_1(t)$.

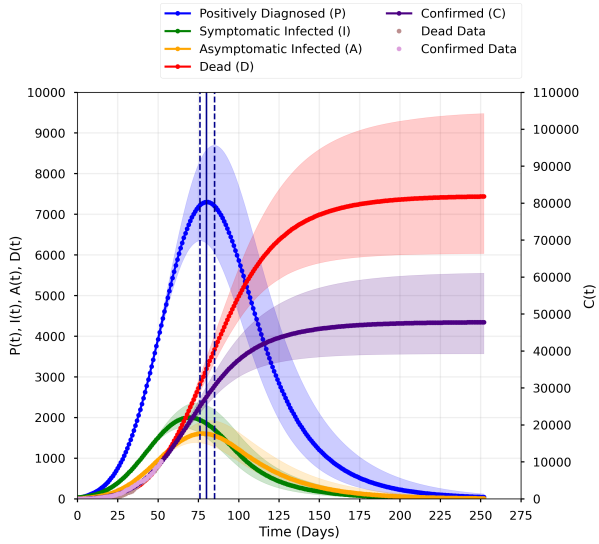


(b) Sensitivity analysis for C and D, $\text{QoI}_2(t)$.

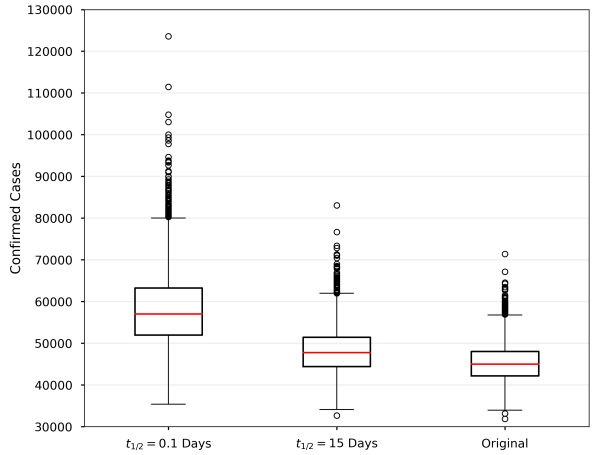
Figure A.14: Temporal changes of the first order sensitivity index of model factors for the RJ scenario (Materials and Methods).



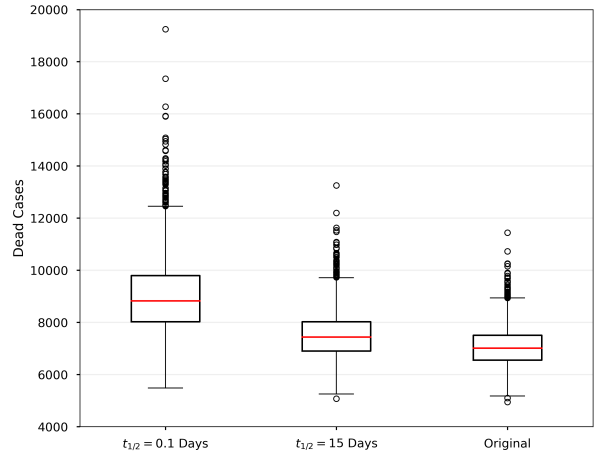
(a) Prediction of P, I, A, D, and C (right y-axis) in RJ considering a sudden release from quarantine ($t_0 = 100$ days and $t_{1/2} = 0.1$ days). The vertical line indicates the curve peak, and the dashed lines display the corresponding 95% CI.



(b) Prediction of P, I, A, D, and C (right y-axis) in RJ considering a gradual release from quarantine ($t_0 = 100$ days and $t_{1/2} = 15$ days). The vertical line indicates the curve peak, and the dashed lines display the corresponding 95% CI.



(c) Box plot for C in RJ at the end of simulation.



(d) Box plot for D in RJ at the end of simulation.

Figure A.15: Model forecasts in RJ considering sudden and gradual releases from quarantine ($t_0 = 100$ days and $t_{1/2} = 0.1$ and 15 days). The original (baseline) scenario of the main text is also presented to ease comparison.

References

[1] World Health Organization, Coronavirus disease 2019 (COVID-19): situation report, 126, www.who.int/emergencies/diseases/novel-coronavirus-2019 (Accessed May 26, 2020).

[2] M. Mesel-Lemoine, J. Millet, P. O. Vidalain, H. Law, A. Vabret, V. Lorin, N. Escriou, M. L. Albert, B. Nal, F. Tangy, A human coronavirus responsible for the common cold massively kills dendritic cells but not monocytes, Journal of Virology 86 (14) (2012) 7577–7587. doi:10.1128/JVI.00269-12.

[3] Worldometer, Coronavirus, https://www.worldometers.info/coronavirus/ (Accessed April 20, 2020).

[4] J. R. Koo, A. R. Cook, M. Park, Y. Sun, H. Sun, J. T. Lim, C. Tam, B. L. Dickens, Interventions to mitigate early spread of SARS-CoV-2 in Singapore: a modelling study, The Lancet Infectious Diseases.doi:10.1016/S1473-3099(20)30162-6.

[5] C. T. Bauch, J. O. Lloyd-Smith, M. P. Coffee, A. P. Galvani, Dynamically modeling SARS and other newly emerging respiratory illnesses: past, present, and future, Epidemiology 16 (6) (2005) 791–801. doi:10.1097/01.ede.0000181633.80269.4c.

[6] COVID-19 Brazil, Análise de subnotificação (in Portuguese), https://ciis.fmrp.usp.br/covid19/analise-subnotificacao/ (Accessed April 20, 2020).

[7] S. B. Bastos, D. O. Cajueiro, Modeling and forecasting the early evolution of the COVID-19 pandemic in Brazil, arXiv e-prints (2020) arXiv:2003.14288.

[8] R. A. Schulz, C. H. Coimbra-Araújo, S. W. S. Costiche, COVID-19: A model for studying the evolution of contamination in Brazil, arXiv e-prints (2020) arXiv:2003.13932.

[9] Decreto 46980 de 19/03/2020, Medidas de enfrentamento, Governo RJ (in Portuguese) (Accessed April 20, 2020).

[10] K. Prem, Y. Liu, T. W. Russell, A. J. Kucharski, R. M. Eggo, N. Davies, et al., The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study, The Lancet Public Health 5 (5) (2020) E261–E270. doi:10.1016/S2468-2667(20)30073-6.

[11] B. F. Maier, D. Brockmann, Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China, Science 368 (6492) (2020) 742–746. doi:10.1126/science.abb4557.

[12] P. Van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, Mathematical Biosciences 180 (1-2) (2002) 29–48. doi:10.1016/S0025-5564(02)00108-6.

[13] N. Crokidakis, COVID-19 spreading in Rio de Janeiro, Brazil: do the policies of social isolation really work?, medRxiv (2020) medRxiv:2020.04.27.20081737doi:10.1101/2020.04.27.20081737.

[14] S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period, Science 368 (6493) (2020) 860–868. doi:10.1126/science.abb5793.

[15] J. Salvatier, T. V. Wiecki, C. Fonnesbeck, Probabilistic programming in Python using PyMC3, PeerJ Computer Science 2 (2016) e55. doi:10.7717/peerj-cs.55.

[16] S. E. Minson, M. Simons, J. L. Beck, Bayesian inversion for finite fault earthquake source models I – theory and algorithm, Geophysical Journal International 194 (3) (2013) 1701–1726. doi:10.1093/gji/ggt180.

[17] D. Volpatto, A. C. M. Resende, L. Anjos, J. V. O. Silva, C. M. Dias, R. C. Almeida, S. M. C. Malta, pydemic: Scripts for BR/RJ social distancing study (May 2020). doi:10.5281/zenodo.3865730.

[18] Brazilian Health Ministry, Coronavirus panel (in Portuguese), https://covid.saude.gov.br/ (Accessed May 26, 2020).

[19] W. Cota, Monitoring the number of COVID-19 cases and deaths in Brazil at municipal and federative units level, SciELOPreprints (2020) SciELOPreprints:362doi:10.1590/scielopreprints.362.

[20] F. Campolongo, J. Cariboni, A. Saltelli, An effective screening design for sensitivity analysis of large models, Environmental modelling & software 22 (10) (2007) 1509–1518. doi:10.1016/j.envsoft.2006.10.004.

[21] J. Herman, W. Usher, SALib: an open-source Python library for sensitivity analysis, Journal of Open Source Software 2 (9) (2017) 97. doi:10.21105/joss.00097.

[22] T. W. Ng, G. Turinici, A. Danchin, A double epidemic model for the SARS propagation, BMC Infectious Diseases 3 (1) (2003) 19. doi:10.1186/1471-2334-3-19.

[23] A. B. Gumel, S. Ruan, T. Day, J. Watmough, F. Brauer, P. Van den Driessche, D. Gabrielson, C. Bowman, M. E. Alexander, S. Ardal, J. Wu, B. M. Sahai, Modelling strategies for controlling SARS outbreaks, Proceedings of the Royal Society of London. Series B: Biological Sciences 271 (1554) (2004) 2223–2232. doi:10.1098/rspb.2004.2800.

[24] Z. Zhang, The outbreak pattern of SARS cases in China as revealed by a mathematical model, Ecological Modelling 204 (3-4) (2007) 420–426. doi:10.1016/j.ecolmodel.2007.01.020.

[25] C. E. Mills, J. M. Robins, M. Lipsitch, Transmissibility of 1918 pandemic influenza, Nature 432 (7019) (2004) 904–906. doi:10.1038/nature03063.

[26] R. J. Hatchett, C. E. Mecher, M. Lipsitch, Public health interventions and epidemic intensity during the 1918 influenza pandemic, Proceedings of the National Academy of Sciences 104 (18) (2007) 7582–7587. doi:10.1073/pnas.0610941104.

[27] D. He, J. Dushoff, T. Day, J. Ma, D. J. D. Earn, Inferring the causes of the three waves of the 1918 influenza pandemic in England and Wales, Proceedings of the Royal Society B: Biological Sciences 280 (1766) (2013) 20131345. doi:10.1098/rspb.2013.1345.

[28] M. M. Saito, S. Imoto, R. Yamaguchi, H. Sato, H. Nakada, M. Kami, S. Miyano, T. Higuchi, Extension and verification of the SEIR model on the 2009 influenza A (H1N1) pandemic in Japan, Mathematical Biosciences 246 (1) (2013) 47–54. doi:10.1016/j.mbs.2013.08.009.

[29] D. K. Mamo, P. R. Koya, Mathematical modeling and simulation study of SEIR disease and data fitting of ebola epidemic spreading in West Africa, Journal of Multidisciplinary Engineering Science and Technology 2 (1) (2015) 106–114.

[30] P. Diaz, P. Constantine, K. Kalmbach, E. Jones, S. Pankavich, A modified SEIR model for the spread of ebola in Western Africa and metrics for resource allocation, Applied Mathematics and Computation 324 (2018) 141–155. doi:10.1016/j.amc.2017.11.039.

[31] L. M. Childs, N. N. Abuelezam, C. Dye, S. Gupta, M. B. Murray, B. G. Williams, C. O. Buckee, Modelling challenges in context: Lessons from malaria, HIV, and tuberculosis, Epidemics 10 (2015) 102–107. doi:10.1016/j.epidem.2015.02.002.

[32] S. Riley, K. Eames, V. Isham, D. Mollison, P. Trapman, Five challenges for spatial epidemic models, Epidemics 10 (2015) 68–71. doi:10.1016/j.epidem.2014.07.001.

[33] J. O. Lloyd-Smith, S. Funk, A. R. McLean, S. Riley, J. L. Wood, Nine challenges in modelling the emergence of novel pathogens, Epidemics 10 (2015) 35–39, challenges in Modelling Infectious Disease Dynamics. doi:10.1016/j.epidem.2014.09.002.

[34] W. Lyra, J. D. Nascimento, J. Belkhiria, L. Almeida, P. P. Chrispim, I. Andrade, COVID-19 pandemics modeling with SEIR(+CAQH), social distancing, and age stratification. the effect of vertical confinement and release in Brazil., medRxiv (2020) medRxiv:2020.04.09.20060053doi:10.1101/2020.04.09.20060053.

[35] H. M. Yang, L. P. L. Junior, A. C. Yang, Modeling the transmission of new coronavirus in São Paulo state, Brazil – Assessing epidemiological impacts of isolating young and elder persons, arXiv e-prints (2020) arXiv:2004.05715.

[36] R. M. Cotta, C. P. Naveira-Cotta, P. Magal, Parametric identification and public health measures influence on the COVID-19 epidemic evolution in Brazil, medRxiv (2020) medRxiv:2020.03.31.2004910doi:10.1101/2020.03.31.20049130.

[37] A. Ducrot, P. Magal, T. Nguyen, G. F. Webb, Identifying the number of unreported cases in SIR epidemic models, Mathematical Medicine and Biology: A Journal of the IMA (2019) 1–19doi:10.1093/imammb/dqz013.

[38] B. Tang, X. Wang, Q. Li, N. L. Bragazzi, S. Tang, Y. Xiao, J. Wu, Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions, Journal of Clinical Medicine 9 (2) (2020) 462. doi:10.3390/jcm9020462.

[39] A. O. Victor, Mathematical predictions for COVID-19 as a global pandemic, medRxiv (2020) medRxiv:2020.03.19.20038794doi:10.1101/2020.03.19.20038794.

[40] L. Peng, W. Yang, D. Zhang, C. Zhuge, L. Hong, Epidemic analysis of COVID-19 in China by dynamical modeling, medRxiv (2020) medRXiv:2020.02.16.20023465doi:10.1101/2020.02.16.20023465.

[41] A. Rovetta, A. S. Bhagavathula, Modelling the epidemiological trend and behavior of COVID-19 in Italy, medRxiv (2020) medRXiv:2020.03.19.20038968doi:10.1101/2020.03.19.20038968.

[42] Q. Lin, S. Zhao, D. Gao, Y. Lou, S. Yang, S. S. Musa, M. H. Wang, Y. Cai, W. Wang, L. Yang, D. He, A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action, International Journal of Infectious Diseases 93 (2020) 211–216. doi:10.1016/j.ijid.2020.02.058.

[43] P. G. T. Walker, C. W., O. W., et al., The global impact of COVID-19 and strategies for mitigation and suppression, Imperial College Report 12 (2020) 1–19. doi:10.25561/77735.

[44] P. V. Savi, M. A. Savi, B. Borges, A mathematical description of the dynamics of coronavirus disease (COVID-19): A case study of Brazil, arXiv e-prints (2020) arXiv:2004.03495.

[45] L. Tarrataca, C. M. Dias, D. B. Haddad, E. F. Arruda, Flattening the curves: on-off lock-down strategies for COVID-19 with an application to Brazi, arXiv e-prints (2020) arXiv:2004.06916.

[46] N. Crokidakis, Data analysis and modeling of the evolution of COVID-19 in Brazil, arXiv e-prints (2020) arXiv:2003.12150.

[47] J. Jia, J. Ding, S. Liu, G. Liao, J. Li, B. Duan, G. Wang, R. Zhang, Modeling the control of COVID-19: impact of policy interventions and meteorological factors, Electronic Journal of Differential Equations 2020 (23) (2020) 1–24.

[48] L. Bao, W. Deng, H. Gao, C. Xiao, J. Liu, J. Xue, Q. Lv, J. Liu, P. Yu, et al., Reinfection could not occur in SARS-CoV-2 infected rhesus macaques, bioRxiv (2020) bioRxiv:2020.03.13.990226doi:10.1101/2020.03.13.990226.

[49] L. Wu, N. Wang, Y. Chang, X. Tian, D. Na, L. Zhang, L. Zheng, T. Lan, L. Wang, G. Liang, Duration of antibody responses after severe acute respiratory syndrome, Emerging Infectious Diseases 13 (10) (2007) 1562–1574. doi:10.3201/eid1310.070576.

[50] L. Petzold, Automatic selection of methods for solving stiff and nonstiff systems of ordinary differential equations, SIAM journal on scientific and statistical computing 4 (1) (1983) 136–148. doi:10.1137/0904010.

[51] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, et al., SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods 17 (2020) 261–272. doi:10.1038/s41592-019-0686-2.

[52] A. C. Hindmarsh, ODEPACK, a systematized collection of ODE solvers, Scientific computing 1 (1983) 55–64.

[53] O. Diekmann, J. A. P. Heesterbeek, M. G. Roberts, The construction of next-generation matrices for compartmental epidemic models, Journal of The Royal Society Interface 7 (2010) 873–885. doi:10.1098/rsif.2009.0386.

[54] A. Tarantola, Inverse problem theory and methods for model parameter estimation, SIAM, Philadelphia, 2005. doi:10.1137/1.9780898717921.

[55] G. A. K. Van Voorn, B. W. Kooi, Combining bifurcation and sensitivity analysis for ecological models, The European Physical Journal Special Topics 226 (9) (2017) 2101–2118. doi:10.1140/epjst/e2017-70030-2.

[56] K. Mizumoto, K. Kagaya, A. Zarebski, G. Chowell, Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020, Eurosurveillance 25 (10) (2020) 2000180. doi:10.2807/1560-7917.

[57] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, J. Lessler, The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application, Annals of internal medicine 172 (9) (2020) 577–582. doi:10.7326/M20-0504.

[58] F. Pan, T. Ye, P. Sun, S. Gui, B. Liang, L. Li, D. Zheng, J. Wang, R. L. Hesketh, L. Yang, et al., Time course of lung changes on chest ct during recovery from 2019 novel coronavirus (COVID-19) pneumonia, Radiology (2020) 200370doi:10.1148/radiol.2020200370.