medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Standard and anomalous second waves in the COVID-19 pandemic

Giovani L. Vasconcelos[1,*], Arthur A. Brum[2], Francisco A. G. Almeida[3], Antônio M. S. Macêdo[2], Gerson C. Duarte-Filho[3], and Raydonal Ospina[4]

[1]Departamento de Física, Universidade Federal do Paraná, 81531-990 Curitiba, Paraná, Brazil
[2]Departamento de Física, Universidade Federal de Pernambuco, 50670-901 Recife, Pernambuco, Brazil
[3]Departamento de Física, Universidade Federal de Sergipe, 49100-000, São Cristóvão, Sergipe, Brazil
[4]Departamento de Estatística, Universidade Federal de Pernambuco, 50740–540, Recife, Pernambuco, Brazil
[*]giovani.vasconcelos@ufpr.br

ABSTRACT

We apply a generalised logistic growth model, with time dependent parameters, to describe the fatality curves of the COVID-19 disease for several countries that exhibit a second wave of infections. The model parameters vary as a function of time according to a logistic function, whose two extreme values, i.e., for early and late times, characterise the first and second waves, respectively. We show that the theoretical curves are in excellent agreement with the empirical data for all cases considered. The model also allows for predictions about the time of occurrence and relative severity of the second wave, in comparison to the first wave. It is shown furthermore that the COVID-19 second waves can be generically classified in two main types, namely, standard and anomalous second waves, according as to whether the second wave starts well after or still during the first wave, respectively. We have also observed that the standard second waves tend, in their majority, to be more severe than the corresponding first wave, whereas for anomalous second waves the opposite occurs.

1 Introduction

One year after the first death by the novel coronavirus (SARS-CoV-2), on January 11, 2020, in Wuhan, China, the direst predictions about the danger and severity of the ensuing pandemic have been confirmed. As of this writing, more than 100 millions of cases of infection by the SARS-CoV-2 virus haven been identified worldwide and over 2.2 millions of deaths have been attributed to the disease (COVID-19) caused by the virus[1,2]. The response strategies to counter the propagation of the virus have varied widely from country to country, and even within countries. Notwithstanding their different approaches to fight the COVID-19 pandemic, a great deal of countries have suffered a significant loss of life to the virus. The widespread severity of the pandemic is reflected in the fact that, as of this writing, 54 countries have registered more than 50 deaths per 100,000 inhabitants, with half of them showing a death rate above 100 deaths per 100,000 inhabitants[3]. A particularly interesting but troublesome development in the course of the pandemic is the fact that many countries were able to control somewhat the spread of the disease during the first few months after its onset, only to see a subsequent increase in the rate of infections and deaths. This resurgence of the COVID-19 epidemic, commonly referred to as a second wave of infections, has in some cases been even more severe that the so-called first wave. Hence it is a topic of special relevance, from both the mathematical modelling viewpoint and the public health perspective.

In the context of the COVID-19 epidemic, a second wave of infections can, broadly speaking, appear via two main different dynamics. First, in a standard or textbook second wave, the resurgence of infections appears after the epidemic curves for the cumulative number of cases and deaths have reached a near-

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

plateau. This means that the daily numbers of new infections and deaths have decreased substantially—and remains low for a somewhat prolonged period of time—, before they surge again. Correspondingly, the daily curves display two well-defined sharp "peaks," separated by a rather shallow "valley." In other words, in a standard second wave there is a clear, distinct separation between the first and the second waves of infections. There are other situations, however, where the epidemic curve undergoes a strong re-acceleration regime even before the daily rates of infections and death have been significantly reduced, indicating that a second wave starts before the first wave has subdued. Such an "anomalous" second-wave effect shows up in the respective cumulative curve as a rapid change in the trend of the growth profile, i.e., from deceleration to acceleration, even though no plateau-like regime had yet been reached. Thus, in an anomalous second wave there is a transition period where the first and the second waves can be said to "coexist," as represented by the fact that the two peaks in the daily curve are separated by a relatively high valley.

Locating and quantifying second-wave effects in a given epidemic curve, beyond simple visual inspection, is an important but not trivial task. This requires, for instance, a mathematical or computational model that is able to efficiently describe the complex growth profiles that arise in the cumulative epidemic curves and from which one can estimate the location and intensity of each wave's peak in the daily curves. A standard way to investigate multiple wave effects is to start with a basic epidemiological model and then allow its parameters to vary in time to reflect the occurrence of secondary waves of infections[4,5]. Understanding the evolution of possible multiple waves of infections is also, of course, relevant for public health officials, as it may help them to develop better strategies to fight the propagation of the virus. In response to the widespread occurrence of second-wave effects in the COVID-19 epidemic in many countries around the world, there is now a growing body of literature on the subject[6–18]. In such studies, compartmental models[10,11,14] and agent based models[19,20] are typically the models of choice, although models based on artificial intelligence algorithms have also been used[15–18].

In this paper, we depart from previous approaches and propose to model second-wave effects in the COVID-19 epidemic in terms of a generalised logistic model with time-dependent parameters. More specifically, we consider an extension of the so-called beta logistic model (BLM)[21], where we assume that each parameter of the model (see below) is allowed to vary continuously and smoothly in time between two well defined values, representing the first and second waves of the epidemic dynamics, respectively. We apply the model to study the fatality curves of COVID-19, as represented by the cumulative number of deaths as a function of time, for several selected countries that display second wave effects, namely: Australia, Austria, Brazil, Germany, Iran, Italy, Japan, Morocco, Serbia, and US. We show that the generalised BLM describes very well the mortality curves of all selected countries. Furthermore, from the theoretical curves we are able to quantify the transition times between the first and second waves as well as the relative intensity of the second wave. In particular, we have observed that standard second waves, in the sense defined above, tend to be more severe than the corresponding first waves; whereas the opposite happens for anomalous second waves.

2 Data

Here we focus exclusively on mortality data from COVID-19, rather than on the number of infection cases, because it is difficult to estimate the actual number of infected people by the SARS-CoV-2, since the confirmed cases represent only an unknown fraction of the total number of infections. In this scenario, the number of deaths attributed to COVID-19 is a somewhat more reliable measure to describe the dynamics of the epidemic[22].

We analysed the COVID-19 fatality curves for ten selected countries, namely: Australia, Austria,

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Brazil, Germany, Iran, Italy, Japan, Morocco, Serbia, and US. As our main aim here is to analyse the second waves of the COVID-19 epidemic in different countries, we have not included mortality data that may be attributed to third or subsequent waves of infections. For example, for Iran, Japan, Morocco, Serbia, and US, we used mortality data up to a maximum date (typically in September or October, 2020) that excludes a third regime of accelerated growth that has been observed in the epidemic curves of these countries. For Australia, we analysed data up to November 6, 2020, because at this point the second wave had nearly reached a plateau, after which no resurgence of COVID-19 deaths has so far been observed. For Austria, Brazil, Germany, and Italy we used recent data (January, 2021), because in these countries the respective second waves are still developing as of the time of this writing.

The data used in this study were obtained from the database made publicly available by the Johns Hopkins University[1,23], which lists in automated fashion the number of the confirmed cases and deaths per country. The specific time periods comprised in the mortality datasets for each country considered in the presented study are given in Sec. 4.

3 Methods

In the present paper we are mainly interested in modelling epidemic curves that display second-wave effects, as indicated for example by the presence of two regimes with strong positive accelerations, corresponding to the first and second waves of infection, respectively. Because our model for such cases is a generalisation of a single-wave model, we shall start by reviewing the basic one-wave model (i.e., with constant parameters), after which the general model with time-dependent parameters is discussed. Subsequently, we also discuss the numerical methods used to analyse the empirical data.

3.1 Single-Wave Model

We model the time evolution of the number of deaths in the epidemic by means of the beta logistic model (BLM), defined by the following ordinary differential equation (ODE)[21,24]:

$$\frac{dC}{dt} = r\,[C(t)]^q \left[1 - \left(\frac{C(t)}{K}\right)^{\alpha}\right]^p, \tag{1}$$

where $C(t)$ is the cumulative number of deaths at time t. We assume for the time being that the model parameters $\{r, q, \alpha, p, K\}$ are all constant in time, in which case they can be interpreted as follows: r is the growth rate at the early stage; q controls the initial growth profile and allows to interpolate from linear growth ($q = 0$) to sub-exponential growth ($q < 1$) to purely exponential growth ($q = 1$); the exponent p controls the late-time growth rate, with $p > 1$ implying a slow-decaying polynomial rate, whereas $p = 1$ yields a fast exponential decay (see below); the exponent α controls the degree of asymmetry with respect to the symmetric S-shape of the logistic curve, which is recovered for $q = p = \alpha = 1$; and, finally, K is the final size of the epidemic, meaning that $C(t) = K$ for $t \to \infty$. Equation (1) must be supplemented with the initial condition

$$C(0) = C_0, \tag{2}$$

for some given value of C_0.

The BLM described in (1) is one of the most general growth models, from which many other known models emerge as special cases[21,24]. For instance, for $q = p = \alpha = 1$ we recover the standard logistic model, as already mentioned. In addition, for $q = p = 1$ we obtain the Richards growth model[25], with the case $p = 1$ corresponding to the so-called generalised Richards model[26]; while setting $\alpha = 1$ in (1) yields the Blumberg's equation[27]; for other special cases see[24].

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

In the case where the parameters $\{r, q, \alpha, p, K\}$ are constant, the BLM admits an analytic solution[21] in the following implicit form:

$$t = f(C) - f(C_0), \tag{3}$$

where

$$f(C) = \frac{C^{1-q}}{r(1-q)}\, {}_2F_1\left(p, \frac{1-q}{\alpha}; 1 + \frac{1-q}{\alpha}; \left(\frac{C}{K}\right)^{\alpha}\right), \tag{4}$$

with ${}_2F_1(a,b;c;x)$ being the Gauss hypergeometric function. Equation (3) describes a sigmoidal curve, whose only inflection point is located at the time t_c obtained by substituting the value $C_c = K[q/(q+\alpha p)]^{1/\alpha}$ in (3). The small- and large-times asymptotic behaviour of the growth profile $C(t)$ are as follows:

$$C(t) \approx \begin{cases} At^{\mu}, & \text{for} \quad t \ll t_c, \\ K - \dfrac{B}{t^{\nu}}, & \text{for} \quad t \gg t_c, \end{cases} \tag{5}$$

where $\mu = 1/(1-q)$, $A = [r(1-q)]^{1/(1-q)}$, $\nu = 1/(p-1)$, and $B = [K^{p-q}/(p-1)r\alpha^p]^{1/(p-1)}$. (For $q \to 1$ and $p \to 1$, one obtains exponential growth and exponential decay, respectively.)

Growth models have the mathematical advantage that they often admit analytical solutions, as we have shown above for the BLM, which is a very useful property when fitting models to empirical data[28]. Furthermore, it is worth pointing out that there is an intrinsic connection between growth models and mechanistic epidemic models of the Susceptible-Infected-Recovered (SIR) class of models. For instance, it is possible to construct a map between the Richards growth model and SIR-type models[29,30]. Thus, when properly applied and interpreted, growth models are useful tools for understanding the spreading dynamics of infectious diseases[21,22,26]. Although in the present paper we restrict our analysis to the BLM, it is important to bear in mind the aforementioned connection between growth models and compartmental models. We therefore thought it was worth including here a brief discussion about a map between the BLM and a generalised SIRD model.

3.2 SIRD Model with Power-Law Behaviour

As mentioned above, it is possible to put the BLM in correspondence with a SIRD-like model, but in this case the target SIRD-type model has to be modified by the inclusion of a power law in the incidence term, owing to the power-law behaviour exhibited by the BLM, as shown below.

We start by recalling the standard Susceptible (S)-Infected (I)-Recovered (R)-Deceased (D) epidemiological model[31,32]

$$\frac{dS(t)}{dt} = -\frac{\beta S(t) I(t)}{N} \tag{6}$$

$$\frac{dI(t)}{dt} = \frac{\beta S(t) I(t)}{N} - (\gamma_1 + \gamma_2) I(t) \tag{7}$$

$$\frac{dR(t)}{dt} = \gamma_1 I(t) \tag{8}$$

$$\frac{dD(t)}{dt} = \gamma_2 I(t), \tag{9}$$

where $S(t)$, $I(t)$, $R(t)$, and $D(t)$ are the number of individuals at time t in the classes of susceptible, infected, recovered, and dead respectively, while N is the constant total number of individuals in the

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

population, so that $N = S(t) + I(t) + R(t) + D(t)$. The parameters β, γ_1 and γ_2 are the transmission, recovery and death rates respectively. The initial values can be chosen to be $S(0) = S_0$, $I(0) = I_0$, with $S_0 + I_0 = N$, and $R(0) = 0 = D(0)$.

We then consider a modified SIRD model, where in (6) and (7) we replace N with only the partial population in the S and I compartments[29]. Furthermore, we follow Refs.[33,34] and replace the term $I(t)$ on the right-hand side of all equations above by $[I(t)]^p$, to obtain

$$\frac{dS}{dt} = -\frac{\beta S(t)}{I(t)+S(t)}[I(t)]^p, \tag{10}$$

$$\frac{dI}{dt} = \frac{\beta S(t)}{I(t)+S(t)}[I(t)]^p - (\gamma_1+\gamma_2)[I(t)]^p, \tag{11}$$

$$\frac{dR}{dt} = \gamma_1[I(t)]^p, \tag{12}$$

$$\frac{dD}{dt} = \gamma_2[I(t)]^p. \tag{13}$$

Although this model is still not general enough to accommodate all the phenomenology of the intervention biased dynamics of the COVID-19 epidemics, it does nonetheless exhibit subexponential behaviour for both short and large time scales in all compartments. In order to show this, we define $y(t) = S(t) + I(t)$ and divide (11) by (10) to obtain

$$\frac{dy}{y} = \frac{1}{R_0}\frac{dS}{S}, \tag{14}$$

where $R_0 = \beta/(\gamma_1+\gamma_2)$. Integrating both sides of (14), and inserting the result into (10), yields an equation of the BLM type:

$$\frac{dS}{dt} = -r[S(t)]^q\left[1-\left(\frac{S(t)}{K}\right)^{\alpha}\right]^p, \tag{15}$$

where $r, K > 0$, $q = 1 + (p-1)/R_0$ and $\alpha = 1 - 1/R_0$. It follows from (15), in comparison with model (1), that $S(t)$ exhibits power-law regimes (for both early and large times) akin to those described by the BLM. Furthermore, it is easy to see that all other compartments, $I(t)$, $R(t)$, and $D(t)$, inherit from $S(t)$ the power-law behaviour, even though their respective equations of motion are not of the BLM-type. Note that the parameters q, p, and α in (15) are not all independent of one another—in fact, the conditions $q < 1$ and $p > 1$ are mutually incompatible. Nonetheless, the preceding qualitative argument shows that the power-law dynamics of the sort predicted by the BLM can in principle be accommodated by compartmental models. We are currently carrying out further research to establish a more complete map between the BLM and a generalised SIRD model where the exponents q, p, and α are all independent of each other.

The BLM with constant parameters has been show to describe remarkably well the first wave of the COVID-19 epidemic for several countries in Europe and North America[21]. However, after the resurgence of the COVID-19 epidemic in many countries (most notably after the Summer of 2020 in the Northern hemisphere), their respective epidemic curves started to exhibit more complex patterns that cannot be captured by the standard BLM. In the next subsection we introduce a generalised version of the BLM with time-dependent parameters, which is much more adequate to describe growth processes with two or more distinct growth phases, corresponding to different waves of infection.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

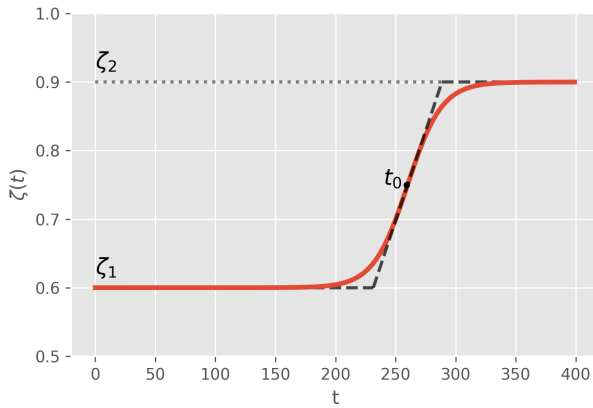

Figure 1. Time dependence of the generic parameter $\zeta(t)$ of the two-wave model, as defined by the logistic function given in (17). The dashed line represents the linear approximation to the logistic function, where the inclined straight line meets the upper and lower horizontal lines at the points $t_0 \pm 2/\rho$, respectively.

3.3 Two-Wave Model

Our two-wave model is still described by the ODE given in (1), but now we assume that *all parameters depend on time*, that is, $r = r(t)$, $q = q(t)$, $\alpha = \alpha(t)$, $p = p(t)$, and $K = K(t)$. To capture the two distinct growth regimes ("waves"), we propose that these parameters, here generically represented by the symbol $\zeta(t)$, obey the following logistic-like equation

$$\frac{d\zeta}{dt} = \rho(\zeta - \zeta_1)\left(1 - \frac{\zeta}{\zeta_2}\right), \tag{16}$$

whose solution, with the condition $\zeta(t_0) = (\zeta_1 + \zeta_2)/2$, is of the following form:

$$\zeta(t) = \zeta_1 + \frac{(\zeta_2 - \zeta_1)}{2}\left[1 + \tanh\left(\frac{\rho(t - t_0)}{2}\right)\right], \tag{17}$$

where $\zeta(t)$ stands for any of the parameters r, α, q, p, and K, with ζ_1 and ζ_2 representing the corresponding parameter values for the first and second waves, respectively. A schematic of the generic parameter $\zeta(t)$, as defined in (17), is shown in Fig. 1. The parameter t_0 determines the transition time between the first and second wave; whereas the parameter ρ characterises how rapid this transition is, so that the larger the parameter ρ the quicker the transition towards the second-wave regime. Note that the characteristic time scale t_0 and the corresponding transition rate ρ are the same for all parameters. This is justified because an overall change in the epidemic dynamics, brought about, say, by a relaxation of control measures or by a change in the population behaviour (or both), is expected to affect simultaneously all epidemiological parameters, which are in turn described in an effective manner by the growth model parameters[22].

It is worth pointing out, however, that although the parameter t_0 in (17) sets the time scale for the transition between the first and second waves, it does not, in itself, represent the time where second-wave effects begin, as the onset of the second wave also depends on the parameter ρ. This can been seen more clearly in Fig. 1, where we also indicated a piecewise linear approximation (dashed line) to the function (solid line) described by (17). Under this linear approximation, one sees that a better estimate for the onset of the second wave is given by $t_0 - 2/\rho$. The linear transition region, $t_0 - 2/\rho < t < t_0 + 2/\rho$, thus

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

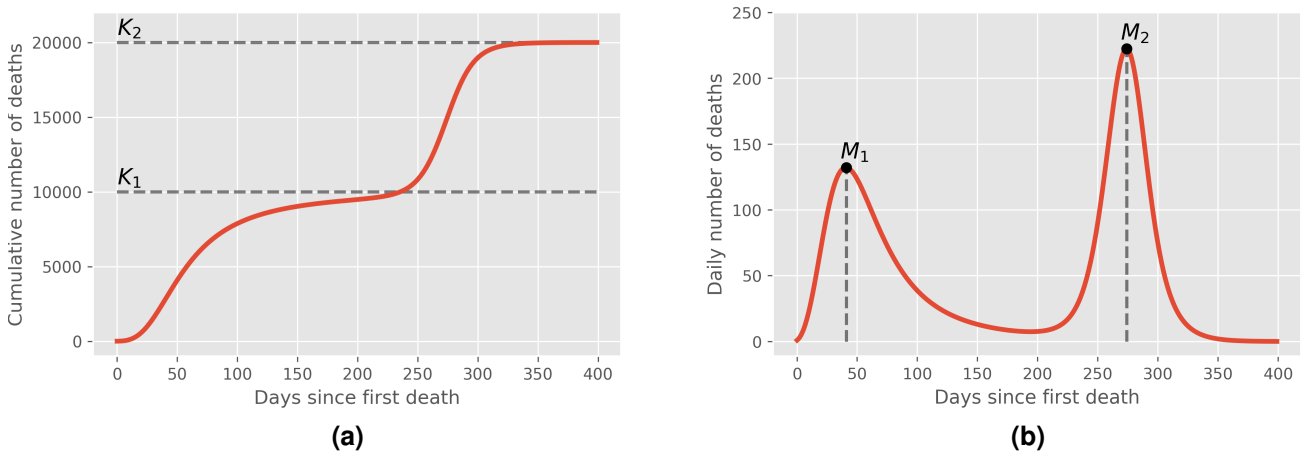

Figure 2. (a) Schematic of an epidemic curve for the cumulative number, $C(t)$, of deaths with two waves of infections. Here K_1 is the plateau value if only the first wave had been present and K_2 is the actual value at the end of the epidemic (final plateau), assuming that no subsequent recrudescence of the epidemic occurs. (b) Time derivative, dC/dt, of the cumulative curve shown in (a), representing the daily number of deaths, where the peaks for the first and second waves are indicated by M_1 and M_2, respectively.

represents the period when new fatalities can, at least from a theoretical viewpoint, be attributed both to the terminal phase of the first wave and the initial phase of the second wave. This superposition effect could be avoided by imposing a discontinuous change of parameters, i.e., taking $\rho \to \infty$, but in this case additional continuity condition on the derivative of $C(t)$ is necessary, as considered , e.g., in Ref.[22]. For numerical purposes, it is more convenient however to describe the transition between the first and second waves with a smooth function, as indicated in (17). Physically, we believe that this smooth transition between different waves of infection is also more reasonable, as a resurgence of infections do not tend to occur suddenly.

As an analytical solution for the theoretical curve $C(t)$ for the BLM time-dependent parameters is no longer possible, one must resort to a numerical integration of the ODE (1), with the parameters $\{r(t), q(t), \alpha(t), p(t), K(t)\}$ described by their respective transition functions of the form given in (17). A schematic of the cumulative curve, $C(t)$, for the two-model upon numerical integration (for an arbitrary choice of parameters) is shown in Fig. 2a. In this figure, the dashed line denoted by K_1 represents the plateau level if only the first wave had been present, whereas the parameter K_2 is the actual final plateau corresponding to the total number of deaths at the end of the epidemic (assuming that subsequent waves of infection do not occur). In Fig. 2b, we show the time derivative, dC/dt, of the cumulative curve shown in Fig. 2a, which corresponds to the daily number of deaths as a function of time. In this figure one clearly sees the distinct "peaks" of the two waves, whose maximum values are indicated by M_1 and M_2, respectively.

At face value, our two-wave model has 12 free parameters, corresponding to the initial and final values for each of the five BLM parameters $\{r, q, \alpha, p, K\}$, together with the parameters t_0 and ρ describing the transition between first and second waves. Blindingly trying to fit a given empirical epidemic curve with a model containing such a large number of parameters is not an efficient procedure, as one is bound to incur into over-fitting issues. Next, we describe a two-step fitting procedure that aims at circumventing, at least partially, these difficulties.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

3.4 Data Analysis

In the first step of our fitting procedure, we give an initial educated guess for the possible location of the transition t_0. We then fit the data up to this time with the one-wave model, as given by its analytical solution (3). At this point, it is important to recall that the parameters r, q, and α in the one-wave model (1) are restricted to certain allowed ranges. For example, the exponent q is limited to the range $0 \leq q \leq 1$, as $q > 1$ would imply a super-exponential growth which is not justified on epidemiological grounds. Furthermore, it is expected by biological reasons (see, e.g., the discussion in Refs.[22,29]) that the asymmetry parameter α should also be within the interval (0,1). Similarly, we restrict the values of r to the range (0,1), as we observed that values of r outside this interval tends to be an indication of possible over-fitting. In other words, we assume here that the restrictions $0 < q \leq 1$, $0 < \alpha \leq 1$, and $0 < r < 1$ are useful empirical criteria to reduce over-fitting.

The second step in our fitting procedure comprises the numerical fit of the entire empirical curve using the two-wave model. In this case, the values obtained for the parameters r_1, q_1, α_1, p_1, and K_1 in the first step described above are used as initial guesses for the respective parameters describing the first wave in the full two-wave model. The initial guesses for the second set of parameters, r_2, q_2, α_2, p_2, and K_2, characterising the second wave as well as for the rate of transition ρ are chosen somewhat arbitrarily within their ranges of definition.

In all numerical fits, for both the single-wave and two-wave models, we employed the Levenberg-Marquardt algorithm to solve the non-linear least square optimisation problem, as implemented in the `lmfit` package for the Python language, which has a built-in routine for estimating the errors of the fitted parameters via the covariance matrix[35]. The results of the fitting procedure are deemed acceptable when the errors in the parameters were smaller than the values themselves estimated for the parameters. (In most cases reported here, however, the errors are much smaller than the maximum allowed tolerance of 100%.) We noticed, however, that in order to control the errors and minimise overfitting issues, it was necessary to apply the restriction $\alpha_2 = 1$ and $p_2 = 1$. This is because these two parameters control the bending and approach towards the second plateau, and since there are generally fewer points in this final portion of the empirical curve (at least for the datasets selected here), estimating these two parameters is more difficult and usually leads to large unacceptable errors. Hence it proved convenient to set $\alpha_2 = p_2 = 1$.

4 Results and Discussion

As our main aim in this paper is to illustrate the application of our two-wave model, we have chosen a representative sample of fatality curves that present a clear, extended second region of accelerated growth, which can be unmistakably associated with a second wave of COVID-19 infections. We have furthermore been careful to avoid including mortality data that may be attributed to possible third or subsequent waves. With these goals in mind, we have analysed the COVID-19 fatality curves for ten selected countries, namely: Australia, Austria, Brazil, Germany, Iran, Italy, Japan, Morocco, Serbia, and US. As discussed before, for some of countries we have included data only up to a maximum earlier date in 2020, so as to exclude third-wave effects; while for others we have include data up to the present time (up to January 2021), because in such cases the second wave of infections is still developing as of the time of writing. (We remark parenthetically that our model can in principle be extended to include any number of subsequent waves, but this requires greater numerical effort, and so this task will not be pursued here; see below for further discussion on this topic.)

In the left panels of Figs. 3-5, we show the cumulative number of deaths (red circles) attributed to COVID-19 for the selected countries, as a function of time counted in days since the first death in each country. Also shown in these figures are the corresponding best fits (black solid curves) by the two-wave

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

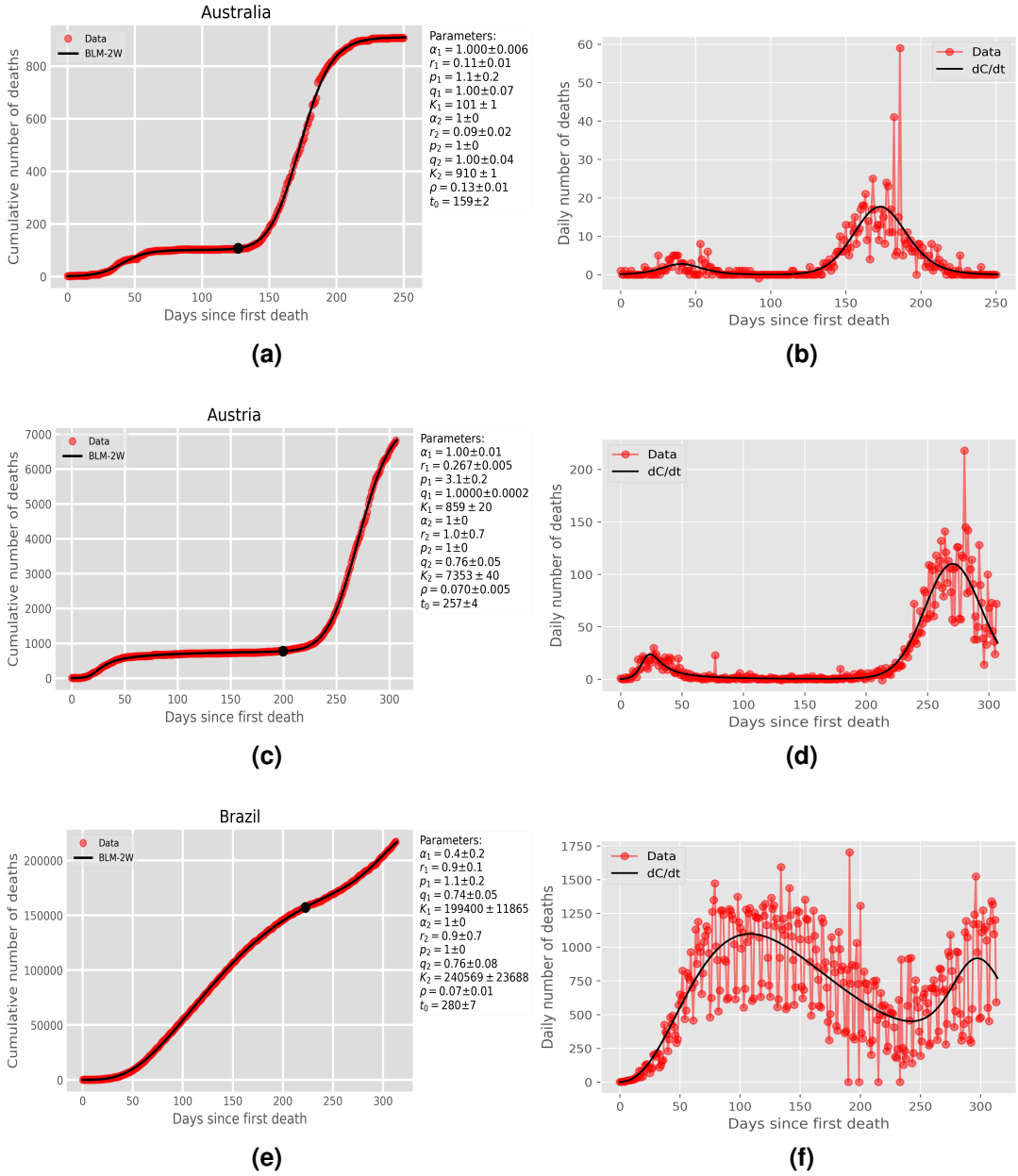

Figure 3. Left panels: Cumulative number of deaths (red circles) attributed to COVID-19 for (a) Australia up to 06/11/2020, (c) Austria up to 12/01/2021, and (e) Brazil up to 24/01/2021. The solid curves are the best fits by the second-wave model, where the black dot in each curve represents the time, $t_0^{(2)} = t_0 - 4/\rho$, that separates the first and second waves. Right panels: Daily number of deaths for the same countries as in the corresponding left panels, where the empirical data are indicated by red circles and the solid curve represents the time derivative of the respective theoretical curve in the left panels.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

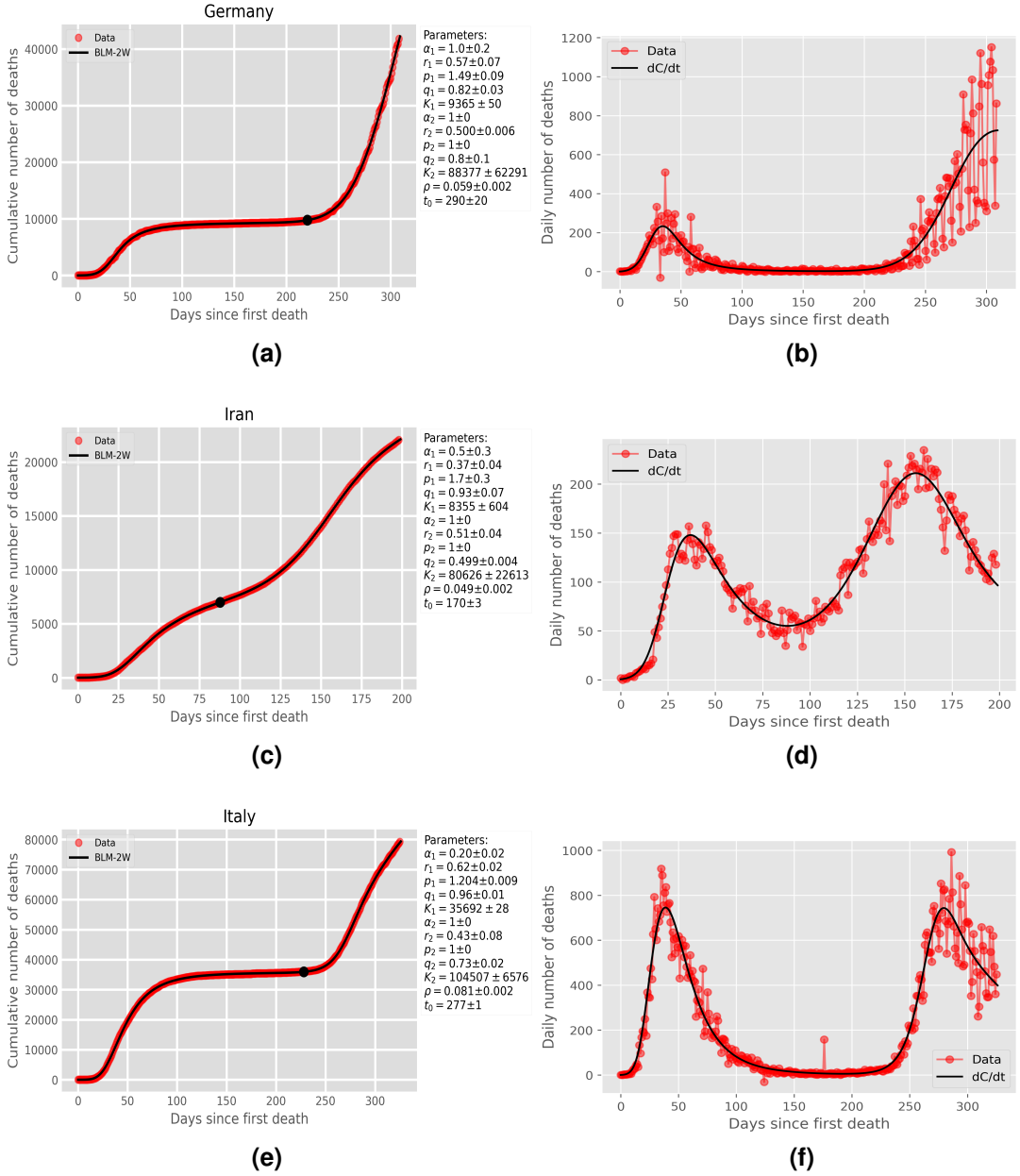

Figure 4. Same as in Fig. 3 for (a) Germany up to 11/01/2021, (c) Iran up to 04/09/2020, and (e) Italy up to 11/01/2021.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

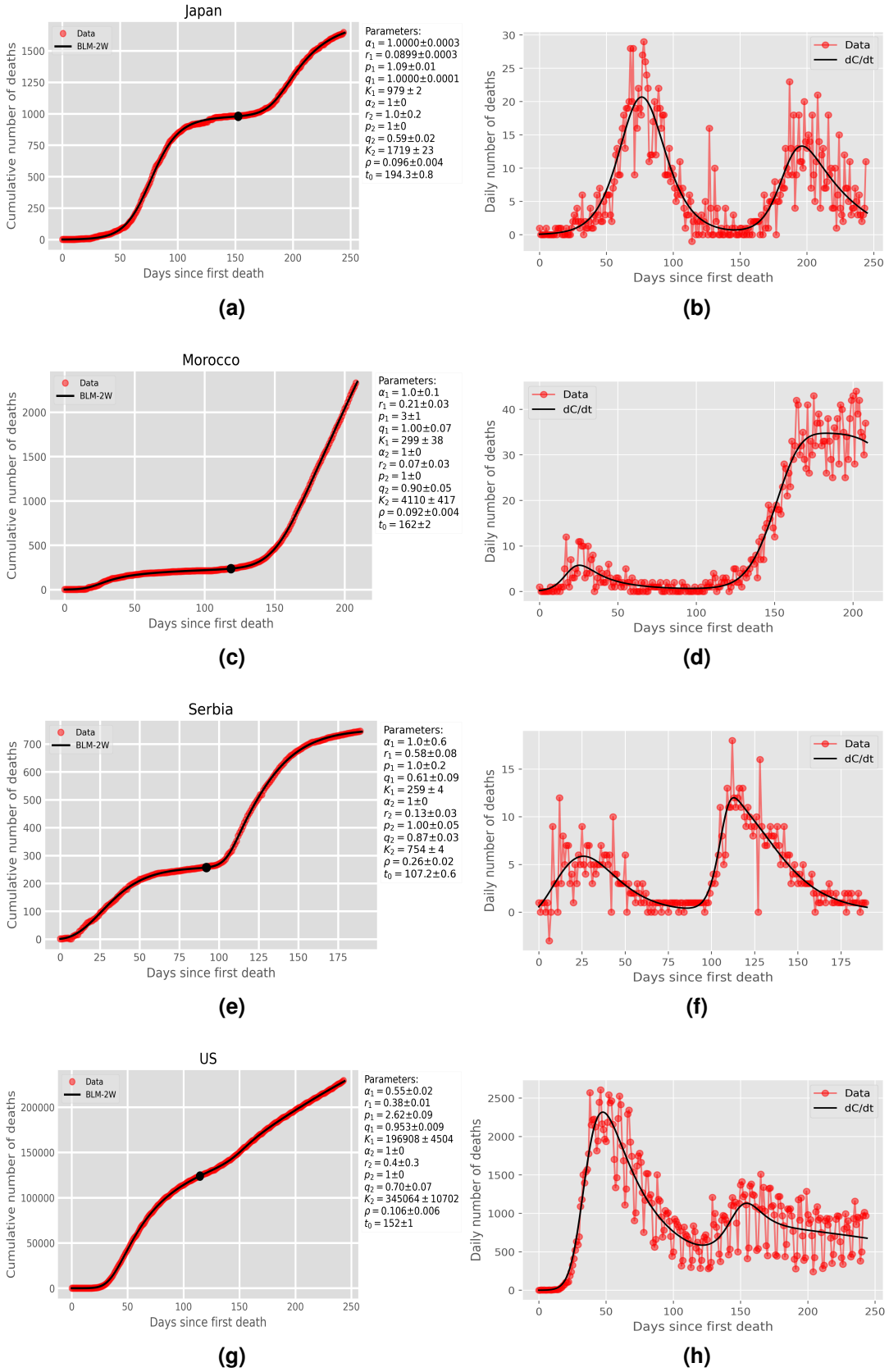

Figure 5. Same as in Fig. 3 for (a) Japan up to 14/10/20, (c) Morocco up to 04/10/2020, (e) Serbia up to 25/09/2020, and (g) US up to 29/10/2020.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

model given in (1) and (17). One sees from these figures that the theoretical curves describe remarkably well the empirical data for all cases considered. The respective best-fit parameters are shown in the legend box of each graph in Figs. 3-5, and the specific dates considered for each country are listed in the captions of the respective figures.

In the right panels of Figs. 3-5 we show the daily numbers of deaths for the selected countries, where the red circles represent the empirical data and the black solid curves correspond to the time derivative of the theoretical curve $C(t)$ predicted by the two-wave model, as obtained from the fits shown in the left panels of the respective figures. One sees that the theoretical daily curves are also in very good agreement with the empirical data. In particular, the model predicts remarkably well the location and general shape of both peaks (in the daily empirical curves) associated with the first and second waves, respectively. It is worth emphasizing that the numerical fits are performed only for the cumulative curves, so that the good agreement between the theoretical daily curves and the empirical daily data represents a further consistency check of the model.

As already discussed in the Introduction, a second wave of infections can, broadly speaking, take place in two main different ways. First, a "standard" second wave pattern can be said to occur when the epidemic curve re-embarks on a rapid acceleration regime after the first wave of infections had nearly 'died out.' This means that the cumulative curve has reached a near-plateau, before it surges up again. Examples of standard second waves can be seen in Figs. 3-5 for the cases of Australia, Austria, Germany, Italy, Morocco, and Serbia. Note that the two peaks in the respective daily curves are well separated by a shallow valley, corresponding to the intervening plateau between the two waves in the cumulative curve. In another possible scenario, an "anomalous" second wave can develop well before the first wave has significantly subsided, causing the cumulative curve to change trend at some point in time (before it reaches a plateau) and re-accelerate again. In such cases, there is a sort of "superposition" of the two effects in the sense that a second wave-like surge appears when the daily deaths (of the first wave) are still relatively high. This implies that the two peaks in the daily cases are not quite far apart, with a relatively high valley between them. Examples of this situation can be seen in Figs. 3-5 for the epidemic curves of Brazil, Iran, Japan, and US.

We have seen from Figs. 3-5 that our two-wave model is capable of describing very well the epidemic curves, over their entire range, for both types of second waves described above. Furthermore, our model has the advantage of being able to provide an estimation for the time of occurrence and intensity of the second wave, as discussed below.

Although a visual inspection of an epidemic curve can easily reveal when a second wave of infection is likely to be present, estimating more precisely when such resurgence actually begins is not an obvious task. As we have argued in Sec. 3.3, the time $t_0 - 2/\rho$ provides an upper bound for the onset of the second wave. This estimate can be further improved by noticing that the width of the transition region between the first and second waves if of the order of $4/\rho$, as illustrated in Fig. 1. Thus, we shall take the time $t_0^{(2)} = t_0 - 4/\rho$ as an estimate for the time when the second-wave effects begin to become important. The corresponding values of $t_0^{(2)}$ for each of the selected countries are indicated in Figs. 3-5 by a black dot on the respective cumulative curve, where one sees that this parameter does indeed correspond to a good separating point between the first and second waves. Next, we now turn our attention to discussing possible measures for the intensity of the second wave.

We recall that the parameter K_1 represents an estimate of the total number of deaths if only the first wave had been present. Similarly, the parameter K_2 is the model prediction for the total number of deaths at the end of the second wave of infections, assuming that a third wave does not occur. Thus, the difference $\Delta K = K_2 - K_1$ can be used to estimate the 'excess of deaths' attributed to the second wave of infections.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 1. Parameters estimating the intensity of the second wave.

Country	K_1	K_2	γ
Australia	101	910	6.00
Austria	859	7353	4.58
Brazil	199400	240569	0.84
Germany	9365	88377	3.11
Iran	8355	80626	1.43
Italy	35692	104507	1.00
Japan	979	1719	0.62
Morocco	299	4110	5.83
Serbia	259	754	2.00
US	196908	345064	0.49

Alternatively, ΔK can be interpreted, at least theoretically speaking, as the number of lives that could have been saved if a given country had been able to avoid its second wave of COVID-19. In Table 1 we show the values of K_1 and K_2 for the 10 countries analysed here. One striking feature of Table 1 is the fact that several countries were very efficient in controlling the first wave of COVID-19, only to see a very high second wave. For example, the excess of deaths, ΔK, attributed to the second wave is nearly 13 times greater than K_1 for Morocco, while this ratio is approximately 9 times for Iran and more than 8 times for Germany, which indicates a heavy toll of lives due to the second COVID-19 wave in comparison to that of the first wave.

It should however be noted that in all selected countries, except perhaps for Australia, the epidemic has not yet (as of now) approached its terminal phase, so that an interpretation of the parameter K_2 as an estimate of the severity of the epidemic requires certain caution. On the one hand, for countries that have experienced a third wave of infections, such as Iran, Japan, Morocco, Serbia, and US, the value of K_2 clearly represents an underestimate of the total number of deaths at the end the epidemic. On the other hand, for countries where the second wave is still developing, such as Austria, Brazil, and Germany, the errors in the estimated values of K_2 tend naturally to be higher, since there are fewer points in the final portion of the curve to allow a more precise estimate of this parameter. Nonetheless, a comparison between K_2 and K_1 is still a valid measure of the severity of the second wave of COVID-19 (relative to the first wave) for all cases above.

An alternate (and possibly more accurate) measure of the intensity of the second wave, relative to the first wave, can be obtained by considering the corresponding peaks in the daily curve. To this end, we introduce the following measure:

$$\gamma = \frac{M_2}{M_1}, \tag{18}$$

where M_1 and M_2 are the maximum values (peaks) of the theoretical daily curves for the first and second waves, respectively. More precisely, we define

$$M_i = \dot{C}(t_i), \quad i = 1, 2, \tag{19}$$

where t_i is the location of the peaks, i.e., $\ddot{C}(t_i) = 0$ and $\dddot{C}(t_i) < 0$, with dots denoting time derivative.

The values of γ for the selected countries are also shown in Table 1. Among the countries that have experienced a standard second wave (Australia, Austria, Italy, Germany, Morocco, and Serbia), only Italy

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

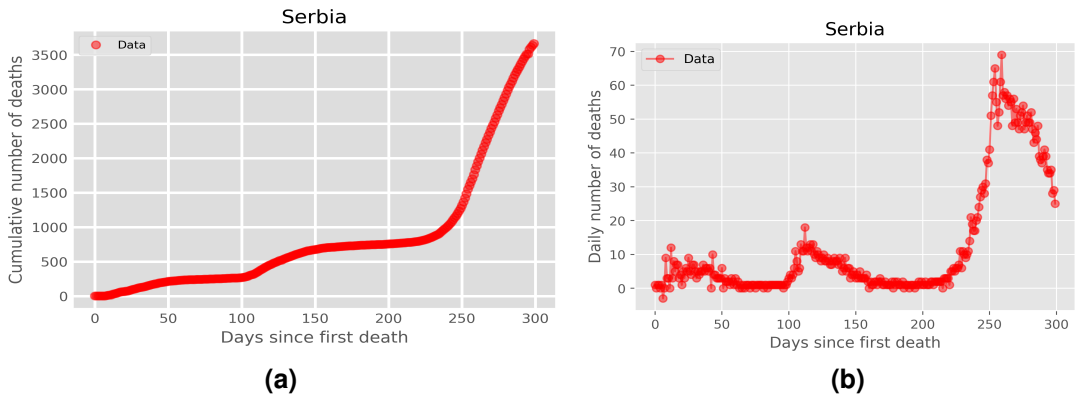

Figure 6. Cumulative (left panel) and daily (right panel) number of deaths attributed to COVID-19 for Serbia up to January 16, 2021, where one can clearly identify three distinct waves of infection.

has a second wave of approximately the same severity as the first wave, as measured by the ratio γ; while for all the others the second wave was considerably more intense. This observation, i.e., that in a standard second-wave scenario the second peak of infections tend to be stronger than the first one, can perhaps be explained as follows: once the first wave was nearly controlled, the enforced or voluntary control measures might have been relaxed (or not adhered by the population with the same zeal), leading to a stronger peak in the second wave. In contradistinction, for countries where the second waves happened in an anomalous way, in the sense explained above, the majority of them (Brazil, Japan, and US) had a relatively less severe second wave, with only Iran showing a more intense second wave. Since by the time of the upsurge of an anomalous second wave, the first wave has not been significantly diminished, it is reasonable to expect that in such cases many of the control measures were likely to be still in place or had only been moderately relaxed, which may explain the lower peak of the second wave. In other words, as health officials and the population in general were still facing a relatively high rate of infections owing to the first wave, it could have been easier in this case to avoid a stronger follow-up wave.

As we mentioned in Sec. 3.4, for some of the selected countries (e.g., Iran, Japan, Morocco, Serbia, US) a third regime of strong re-acceleration has occurred subsequently to the maximum time considered in this study. An example, for illustrative purposes only, is shown in Fig. 6 for Serbia, where one clear sees three distinct waves of infections of increasing intensity. In the present analysis we have not included mortality data that may be attributed to third-wave effects, as mentioned above. Notwithstanding, it should be noted that our model could in principle be extended to include such multiple waves by considering time-dependent parameters of the following form:

$$\zeta(t) = \zeta_1 + \sum_{i=1}^{N-1} \frac{(\zeta_{i+1} - \zeta_i)}{2} \left[1 + \tanh\left(\frac{\rho_i (t - t_i)}{2} \right) \right], \tag{20}$$

where N indicates the number of infection waves. As the total number of parameters grows linearly with N, the numerical task of fitting a N-wave model to a given empirical curve becomes quite challenging, specially considering that the total number of points in an empirical epidemic curve is relatively small (typically of the order of a few hundreds). Nonetheless, this interesting prospect deserves further investigation and is topic of ongoing research.

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

5 Conclusion

In this paper, we have studied the dynamics of the second waves the infections by the novel coronavirus. To this end, we have introduced a generalised logistic model with time-dependent parameters to analyse the COVID-19 fatality curves of several countries from six continents. Not only the theoretical curves are in excellent agreement with the empirical data for all cases considered, but they also allow us to infer predictions about the location and severity of the secondary waves. For instance, we have verified that the intensity of a standard second wave (i.e., one that follows after the first wave has nearly subsided) tends, in its majority, to be considerably stronger that the first wave, as measured by ratio between the respective peaks in the daily number of deaths. Contrarily, in the majority of countries that experienced an anomalous second wave (i.e., where the resurgence of infections takes places while the first wave was still developing), the peak of the second wave was less intense than that of the first wave.

We have argued that the behaviour described above might be due to the possible relaxation of control measures by government and health officials or lack of adherence to them by the respective populations. In other words, a sort of "prevention paradox" might be at play here, whereby early success in controlling the first wave of COVID-19 might lead to a false impression that "the worst was behind," thus stimulating a relaxation of voluntary or enforced measures beyond what would be desirable. By the same token, in countries where a recrudescence of infections happened when the daily rates of deaths were still relatively high, so that more stringent control measures were possibly still in place, such an anomalous second wave would tend to be less severe. Of course, the severity of the secondary waves of infections is not only determined by their time of occurrence, as the dynamics of an epidemic at any stage depends considerably on the specific responses adopted by the affected populations. Furthermore, we have analysed only ten representative countries, as our main objective here was to introduce and validate our second-wave model. Nevertheless, we expect that the trends identified here should hold in general.

The results reported in the present paper are relevant both from a mathematical viewpoint, in that they show that additional care is needed when modelling multiple-wave epidemics, and from a practical perspective, for they may help policymakers and health authorities in devising strategies to battle the disease during all of its waves. Here we have restricted our analysis to second-wave effects, but our model can in principle be extended to include multiple waves. This interesting possibility is currently under investigation.

References

1. Johns Hopkins University. Coronavirus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). https://coronavirus.jhu.edu/map.html (2021). Accessed: 2021-01-30.

2. Worldometer. Worldometer - COVID-19 data. https://www.worldometers.info/coronavirus/ (2020). Accessed: 2021-01-30.

3. Johns Hopkins Coronavirus Resource Center. Mortality analyses. https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases (2021). Accessed: 2021-01-30.

4. Brauer, F., Castillo-Chavez, C. & Feng, Z. *Mathematical models in epidemiology*, vol. 32 (Springer, 2019).

5. Doménech-Carbó, Antonio and Doménech-Casasús, Clara. The evolution of COVID-19: A discontinuous approach. *Phys. A* **568**, 125752–125752 (2021).

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

6. Faranda, D. & Alberti, T. Modeling the second wave of COVID-19 infections in France and Italy via a stochastic SEIR model. *Chaos: An Interdiscip. J. Nonlinear Sci.* **30**, 111101 (2020).

7. Liao, Z., Lan, P., Liao, Z., Zhang, Y. & Liu, S. TW-SIR: time-window based SIR for COVID-19 forecasts. *Scientific Reports* **10**, 1–15 (2020).

8. Kaxiras, E. & Neofotistos, G. Multiple epidemic wave model of the COVID-19 pandemic: Modeling study. *J. Med. Internet Res.* **22**, e20912 (2020).

9. Fan, G. *et al.* Decreased case fatality rate of COVID-19 in the second wave: a study in 53 countries or regions. *Transboundary and Emerging Diseases* (2020).

10. Friston, K. J. *et al.* Second waves, social distancing, and the spread of COVID-19 across America. *arXiv preprint arXiv:2004.13017* (2020). 2004.13017.

11. Friston, K. *et al.* Dynamic causal modelling of COVID-19. *Wellcome Open Res.* **5**, DOI: 10.12688/wellcomeopenres.15881.2 (2020).

12. Dimaschko, J. Superspreading as a Regular Factor of the COVID-19 Pandemic: II. Quarantine Measures and the Second Wave. https://doi.org/10.1101/2020.08.14.20174557 (2020). 2020.08.14.20174557.

13. El Aferni, A., Guettari, M. & Tajouri, T. Mathematical model of Boltzmann's sigmoidal equation applicable to the spreading of the coronavirus (COVID-19) waves. *Environ. Sci. Pollut. Res.* 1–9 (2020).

14. Cacciapaglia, G., Cot, C. & Sannino, F. Second wave COVID-19 pandemics in Europe: a temporal playbook. *Scientific Reports* **10**, 1–8 (2020).

15. Tayarani-N, Mohammad-H. Applications of artificial intelligence in battling against covid-19: a literature review. *Chaos, Solitons & Fractals* 110338–110338 (2020).

16. Tat Dat, T. *et al.* Epidemic dynamics via wavelet theory and machine learning with applications to Covid-19. *Biology* **9**, 477 (2020).

17. Vaid, S., McAdie, A., Kremer, R., Khanduja, V. & Bhandari, M. Risk of a second wave of Covid-19 infections: using artificial intelligence to investigate stringency of physical distancing policies in north america. *Int. Orthop.* **44**, 1581–1589 (2020).

18. Syeda, H. B. *et al.* Role of machine learning techniques to tackle the COVID-19 crisis: Systematic review. *Journal of Medical Internet Research* **9**, e23811 (2021).

19. Ferguson, N. M. *et al.* Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College COVID-19 Response Team (2020).

20. Cacciapaglia, G., Cot, C. & Sannino, F. Multiwave pandemic dynamics explained: How to tame the next wave of infectious diseases. https://arxiv.org/abs/2011.12846 (2020). 2011.12846.

21. Vasconcelos, G. L. *et al.* Power law behaviour in the saturation regime of fatality curves of the COVID-19 pandemic. *Sci. Reports* **11**, 4619, DOI: 10.1038/s41598-021-84165-1 (2021).

22. Vasconcelos, G. L. *et al.* Modelling fatality curves of COVID-19 and the effectiveness of intervention strategies. *PeerJ* **8**, e9421, DOI: 10.7717/peerj.9421 (2020).

23. Humanitariam Data Exchange. Novel Coronavirus (COVID-19) Cases Data. https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases (2020). Accessed: 2021-01-30.

24. Tsoularis, A. & Wallace, J. Analysis of logistic growth models. *Math. Biosciences* **179**, 21–55 (2002).

medRxiv preprint doi: https://doi.org/10.1101/2021.01.31.21250867; this version posted March 1, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

25. Richards, F. A flexible growth function for empirical use. *J. experimental Bot.* **10**, 290–301 (1959).

26. Chowell, G. *et al.* Using phenomenological models to characterize transmissibility and forecast patterns and final burden of Zika epidemics. *PLOS Curr. Outbreaks* (2016).

27. Blumberg, A. Logistic growth rate functions. *J. Theor. Biol.* **21**, 42–44 (1968).

28. Chowell, G., Sattenspiel, L., Bansal, S. & Viboud, C. Mathematical models to characterize early epidemic growth: A review. *Phys. of Life Reviews* **18**, 66–97 (2016).

29. Wang, X.-S., Wu, J. & Yang, Y. Richards model revisited: Validation by and application to infection dynamics. *Journal of Theoretical Biology* **313**, 12–19 (2012).

30. Macêdo, A. M. *et al.* A comparative analysis between a SIRD compartmental model and the Richards growth model. *Tend. Mat. Apl. Comput.* , to appear (2021).

31. Anderson, R. M. *The population dynamics of infectious diseases: theory and applications* (Springer, 2013).

32. Diekmann, O., Heesterbeek, H. & Britton, T. *Mathematical tools for understanding infectious disease dynamics*, vol. 7 (Princeton University Press, 2012).

33. Chowell, G. Fitting dynamic models to epidemic outbreaks with quantified uncertainty: a primer for parameter uncertainty, identifiability, and forecasts. *Infect. Dis. Model.* **2**, 379–398 (2017).

34. Tirnakli, U. & Tsallis, C. Epidemiological model with anomalous kinetics: Early stages of the COVID-19 pandemic. *Front. Phys. 8: 613168. doi: 10.3389/fphy* (2020).

35. Newville, M., Stensitzki, T., Allen, D. & Ingargiola, A. Non-linear least-squares minimization and curve-fitting for Python. *Chicago, IL* (2015).