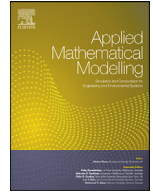
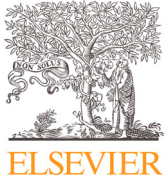




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



# Global analysis of the COVID-19 pandemic using simple epidemiological models



José Enrique Amaro<sup>a</sup>, Jérémie Dudouet<sup>b</sup>, José Nicolás Orce<sup>c,\*</sup>

<sup>a</sup> Departamento de Física Atómica, Molecular y Nuclear and Instituto Carlos I de Física Teórica y Computacional, Universidad de Granada, Granada E-18071, Spain

<sup>b</sup> Univ Lyon, Univ Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, UMR 5822, Villeurbanne F-69622, France

<sup>c</sup> Department of Physics & Astronomy, University of the Western Cape, P/B X17 Bellville ZA-7535, South Africa

## ARTICLE INFO

### Article history:

Received 18 August 2020

Revised 3 October 2020

Accepted 11 October 2020

Available online 22 October 2020

### Keywords:

COVID-19 coronavirus

Death model

Extended SIR model

Monte Carlo Planck model

## ABSTRACT

Several analytical models have been developed in this work to describe the evolution of fatalities arising from coronavirus COVID-19 worldwide. The Death or 'D' model is a simplified version of the well-known SIR (susceptible-infected-recovered) compartment model, which allows for the transmission-dynamics equations to be solved analytically by assuming no recovery during the pandemic. By fitting to available data, the D-model provides a precise way to characterize the exponential and normal phases of the pandemic evolution, and it can be extended to describe additional spatial-time effects such as the release of lockdown measures. More accurate calculations using the extended SIR or ESIR model, which includes recovery, and more sophisticated Monte Carlo grid simulations – also developed in this work – predict similar trends and suggest a common pandemic evolution with universal parameters. The evolution of the COVID-19 pandemic in several countries shows the typical behavior in concord with our model trends, characterized by a rapid increase of death cases followed by a slow decline, typically asymmetric with respect to the pandemic peak. The fact that the D and ESIR models predict similar results – without and with recovery, respectively – indicates that COVID-19 is a highly contagious virus, but that most people become asymptomatic (D model) and eventually recover (ESIR model).

© 2020 Elsevier Inc. All rights reserved.

## 1. Motivation

The SIR (susceptible-infected-recovered) model is widely used as first-order approximation to viral spreading of contagious epidemics [1], mass immunization planning [2,3], marketing, informatics and social networks [4]. Its cornerstone is the so-called “mass-action” principle introduced by Hamer, which assumes that the course of an epidemic depends on the rate of contact between susceptible and infected individuals [5]. This idea was extended to a continuous time framework by Ross in his pioneering work on malaria transmission dynamics [6–8], and finally put into its classic mathematical form by Kermack and McKendrick [9]. The SIR model was further developed by Kendall, who provided a spatial generalization of the Kermack and McKendrick model in a closed population [10] (i.e. neglecting the effects of spatial migration), and Bartlett, who – after investigating the connection between the periodicity of measles epidemics and community size – predicted

\* Corresponding author.

E-mail addresses: [amaro@ugr.es](mailto:amaro@ugr.es) (J. Enrique Amaro), [j.dudouet@ip2i.in2p3.fr](mailto:j.dudouet@ip2i.in2p3.fr) (J. Dudouet), [jnorce@uwc.ac.za](mailto:jnorce@uwc.ac.za) (J. Nicolás Orce).

a traveling wave of infection moving out from the initial source of infection [11,12]. More recent implementations have considered the typical incubation period of the disease and the spatial migration of the population.

The COVID-19 pandemic has ignited the submission of multiple manuscripts in the last weeks. Most statistical distributions used to estimate disease occurrence are of the binomial, Poisson, Gaussian, Fermi or exponential types. Despite their intrinsic differences, these distributions generally lead to similar results, assuming independence and homogeneity of disease risks [13].

In this work, we propose a simple and easy-to-use epidemiological model – the Death or D model [14] – that can be compared with data in order to investigate the evolution of the infection and deviations from the predicted trends. The D model is a simplified version of the SIR model with analytical solutions under the assumption of no recovery – at least during the time of the pandemic. We apply it globally to countries where the infestation of the COVID-19 coronavirus has widespread and caused thousands of deaths [15,16].

Additionally, D-model calculations are benchmarked with more sophisticated and reliable calculations using the extended SIR (ESIR) and Monte Carlo Planck (MCP) models – also developed in this work – which provide similar results, but allow for a more coherent spatial-time disentanglement of the various effects present during a pandemic. A similar ESIR model has recently been proposed by Squillante and collaborators for infected individuals as a function of time, based on the Ising model – which describes ferromagnetism in statistical mechanics – and a Fermi-Dirac distribution [17]. This model also reproduces *a posteriori* the COVID-19 data for infestations in China as well as other pandemics such as Ebola, SARS, and influenza A/H1N1.

The SIR model considers the three possible states of the members of a closed population affected by a contagious disease. It is, therefore, characterized by a system of three coupled non-linear ordinary differential equations [18], which involve three time-dependent functions:

- Susceptible individuals,  $S(t)$ , at risk of becoming infected by the disease.
- Infected individuals,  $I(t)$ .
- Recovered or removed individuals,  $R(t)$ , who were infected and may have developed an immunity system or die.

The SIR model describes well a viral disease, where individuals typically go from the susceptible class  $S$  to the infected class  $I$ , and finally to the removed class  $R$ . Recovered individuals cannot go back to be susceptible or infected classes, as it is, potentially, the case of bacterial infection. The resulting transmission-dynamics system for a closed population is described by

$$\frac{dS}{dt} = -\lambda SI, \quad (1)$$

$$\frac{dI}{dt} = \lambda SI - \beta I, \quad (2)$$

$$\frac{dR}{dt} = \beta I, \quad (3)$$

$$N = S(t) + I(t) + R(t), \quad (4)$$

where  $\lambda > 0$  is the transmission or spreading rate,  $\beta > 0$  is the removal rate and  $N$  is the fixed population size, which implies that the model neglects the effects of spatial migration. Currently, there is no vaccination available for COVID-19, and the only way to reduce the transmission or infection rate  $\lambda$  – which is often referred to as “flattening the curve” – is by implementing strong social distancing and hygiene measures.

The system is reduced to a first-order differential equation, which does not possess an explicit solution, but can be solved numerically. The SIR model can then be parametrized using actual infection data to solve  $I(t)$ , in order to investigate the evolution of the disease. In the D model, we make the drastic assumption of *no recovery* in order to obtain an analytical formula to describe – instead of infestations – the death evolution by COVID-19. This can be useful as a fast method to foresee the global behavior as a first approach, before applying more sophisticated methods. We shall see that the resulting D model describes well enough the data of the current pandemics in different countries.

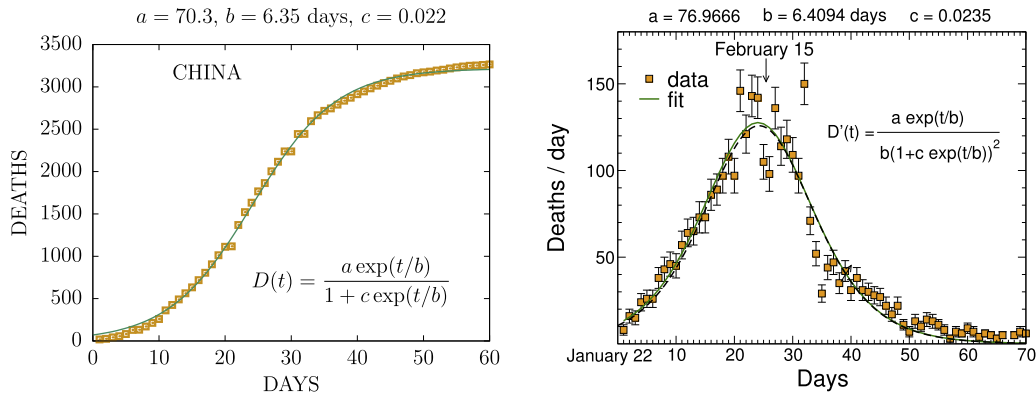
## 2. The death or D model

The main assumption of the D model is the absence of recovery from coronavirus, i.e.  $R(t) = 0$ , at least during the pandemic time interval. This assumption may be reasonable if the spreading time of the pandemic is much faster than the recovery time, i.e.  $\lambda \gg \beta$ . The SIR equations are then reduced to the single equation of the well-known SI model,

$$\frac{dI}{dt} = \lambda(N - I)I(t), \quad (5)$$

which represents the simplest mathematical form of all disease models, where the infection rate is proportional to both the infected,  $I$ , and susceptible individuals  $N - I$ . Eq. (5) is trivially solved by multiplying by  $dt$  and dividing by  $(N - I)I$ ,

$$\frac{dI}{(N - I)I} = \lambda dt, \quad (6)$$



**Fig. 1.** Fits to total (left panel) and daily (right panel) deaths by COVID-19 in China using the  $D(t)$  and  $D'(t)$ , respectively. The dashed curve shows a fit to the daily deaths using the parameters determined to fit the total deaths (top of left panel), which provides similar results to an independent fit (parameters on the top right), given the statistical fluctuations in the daily rates. Data are taken from [19].

where we have defined the constants

$$b = \frac{1}{\lambda N}, \quad C = \frac{I_0}{N}. \quad (7)$$

The parameter  $b$  is the characteristic evolution time of the initial exponential increase of the pandemic. The constant  $C$  is the initial infestation rate with respect to the total population  $N$ . Assuming  $C \ll 1$ ,

$$I(t) = \frac{I_0 e^{(t-t_0)/b}}{1 + C e^{(t-t_0)/b}}. \quad (8)$$

In order to predict the number of deaths in the D model we assume that the number of deaths at some time  $t$  is proportional to the infestation at some former time  $\tau$ , that is,

$$D(t) = \mu I(t - \tau), \quad (9)$$

where  $\mu$  is the death rate, and  $\tau$  is the death time. With this assumption we can finally write the D-model equation as

$$D(t) = \frac{a e^{(t-t_0)/b}}{1 + c e^{(t-t_0)/b}}, \quad (10)$$

where  $a = \mu I_0 e^{-\tau/b}$ ,  $c = C e^{-\tau/b}$ , and  $a/c$  yields the total number of deaths predicted by the model. This is the final equation for the D-model, which presents a similar shape to the well-known Woods–Saxon potential for the nucleons inside the atomic nucleus or the bacterial growth curve. The rest of the parameters,  $\mu$ ,  $\tau$ ,  $I_0$  and  $N$  are embedded in the parameters  $a$ ,  $b$ ,  $c$ , which represent space-time averages and can be fitted to the timely available data. Consequently, the D-function passes into a well-known logistic model, which is described by the Riccati equation, but with different constants (e.g. see Ref. [20] and references therein).

In Fig. 1, we present the fit of the D-model to the COVID-19 death data for China, where its evolution has apparently been controlled and the D function has reached the plateau zone, with few increments over time, or fluctuations that are beyond the model assumptions. This plot shows the duration of the pandemic – about two months to reach the top end of the curve – and the agreement, despite the crude assumptions, between data and the evolution trend described by the D-model. This agreement encourages the application of the D model to other countries in order to investigate the different trends.

### 2.1. Evolution of D-model parameters

In order to get insight into the stability and uncertainty of our predictions, Fig. 2 shows the evolution of  $a$ ,  $b$ , and  $c$  and other model predictions from fits to the daily data in Spain. The meaning of these quantities is explained below:

- The parameter  $a$  is the theoretical number of deaths at the day corresponding to  $t = 0$ . In general, it differs from the experimental value and can be interpreted as the expected value of deaths that day. Note that experimental data may be subject to unknown systematic errors and different counting methods.
- The parameter  $b$ , as mentioned above, is the characteristic evolution time. During the initial exponential behavior, it indicates the number of days for the number of deaths to double. Moreover,  $1/b$  is proportional to the slope of the almost linear behavior in the mid region of the D function. That behavior can be obtained by doing a Taylor expansion around  $t_0 = -b \ln c$  and is given by

$$D(t) \simeq \frac{1}{c} \left( 1 - \frac{1}{2} \ln c \right) + \frac{t}{2bc}. \quad (11)$$

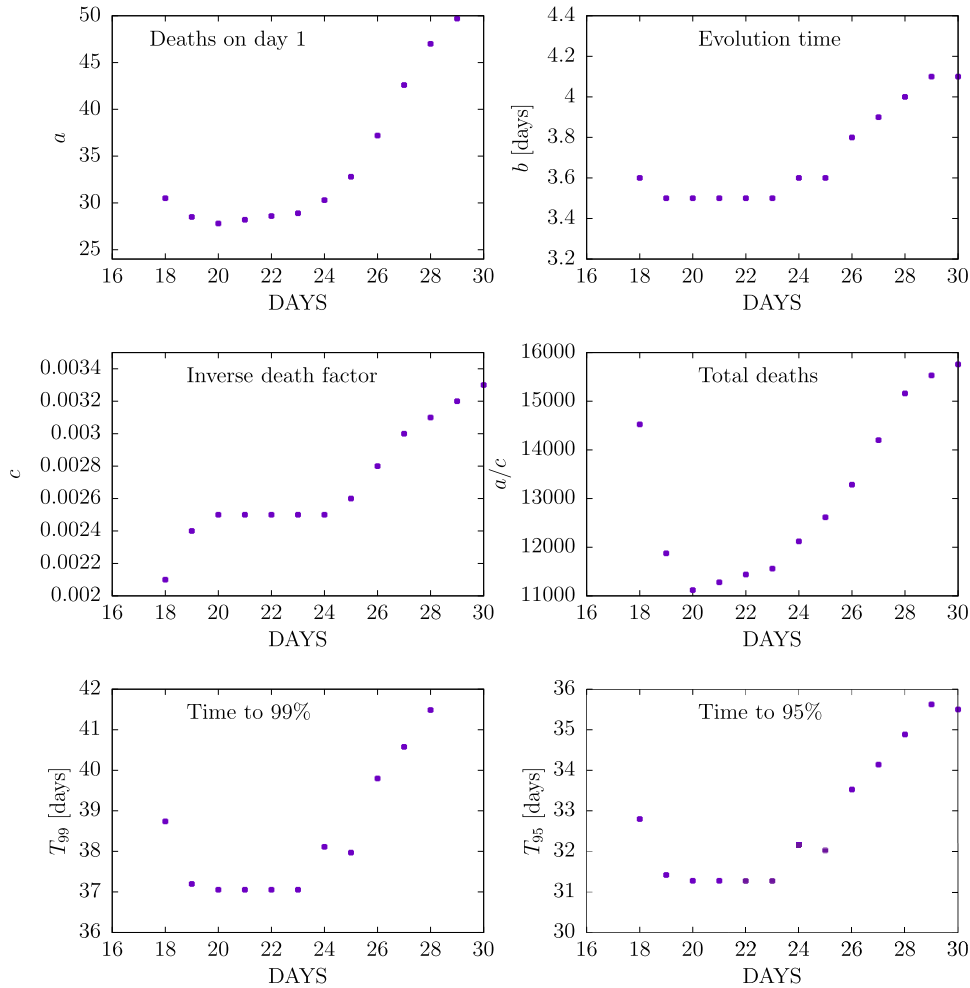


Fig. 2. Evolution of  $a$ ,  $b$  and  $c$  parameters and various predictions of the D-model as a function of time (days).

- The parameter  $c$  is called the *inverse dead factor* because  $D(t \rightarrow \infty) = a/c$  provides the asymptotic or expected total number of deaths.
- The times  $T_{95}$  and  $T_{99}$  correspond to  $D = 0.95D(\infty)$  and  $D = 0.99D(\infty)$ , respectively. These times are obtained by solving the equation  $D(t) = \gamma a/c$ , where  $\gamma = 0.95$  or  $0.99$ . The solution of that equation is

$$t = b \ln \left( \frac{1}{\gamma} \frac{\gamma}{1 - \gamma} \right). \quad (12)$$

Fig. 2 shows the stable trend of the parameters between days 19 and 24 (corresponding to March 27–30), right before reaching the peak of deaths cases, which occurred in Spain around April 1. Such stability validates the D-model predictions during this time. However, a rapid change of the parameters is observed, especially for  $a$ , once the peak is reached, drastically changing the prediction of the number of deaths given by  $a/c$ . This sudden change results in the slowing down of deaths per day and longer time predictions  $T_{95}$  and  $T_{99}$ .

The parameters of the D model correspond to average values over time of the interaction coefficients between individuals, i.e. they are sensitive to an additional external effect on the pandemic evolution. These may include the lockdown effect imposed in Spain in March 14 and other effects such as new sources of infection or a sudden increase of the total susceptible individuals due to social migration and large mass gatherings [21]. It is not possible to identify a specific cause because its effects are blurred by the stochastic evolution of the pandemic, which is why any reliable forecast presents large errors.

## 2.2. The D' model

One can also determine deaths/day rates by applying the first derivative to Eq. (10),

$$D'(t) = \frac{ae^{(t-t_0)/b}}{b(1 + ce^{(t-t_0)/b})^2}, \quad (13)$$

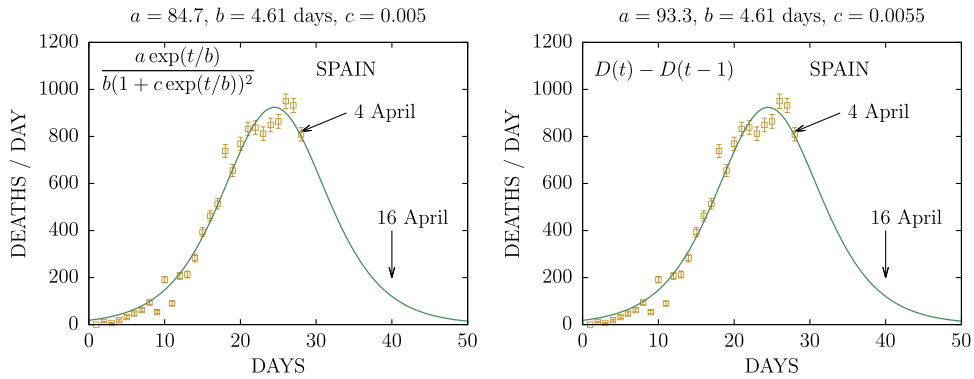


Fig. 3. Predictions of the D model for  $D'(t)$  in Spain according with the data collected up to April 5.

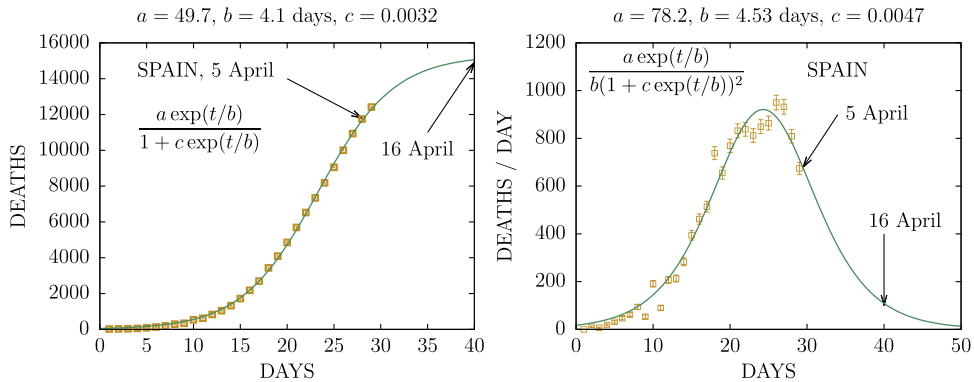


Fig. 4. Comparison of parameters fitted to  $D(t)$  and  $D'(t)$  in Spain according with the data on April 5.

which allows for a determination of the pandemics peak and evolution after its turning point. The  $D$  model describes well the cumulative deaths because the sum of discrete data reduce the fluctuations, in the same way as the integral of a discontinuous function is a continuous function. However, the daily data required for  $D'$  have large fluctuations – both statistical and systematic – which normally gives a slightly different set of parameters when compared with the  $D$  model.

Using the  $D$  model fitted to cumulative deaths allows to compute deaths/day as

$$D(t) - D(t - \Delta t) \simeq D'(t) \Delta t, \quad (14)$$

where  $\Delta t = 1$  day. Fig. 3 shows that Eqs. (13) and (14) yield similar parameters, as the time increment is small enough compared with the time evolution of the  $D(t)$  function. Hence, the first derivative  $D'(t)$  can be used to describe deaths per day. In addition, Fig. 4 shows that the parameters may be different for both  $D$  and  $D'$  functions using cumulative and daily deaths, respectively, as shown for Spain on April 5. It is also important to note that  $b$  is directly proportional to the full width at half maximum (FWHM) of the  $D'(t)$  distribution,

$$FWHM = 2b \ln(3 + 2\sqrt{2}) \approx 3.5 b. \quad (15)$$

As shown below, the  $b$  parameter presents typical values between 4 and 10 for most countries undergoing the initial exponential phase, which yields a minimum and maximum time of 14 and 35 days, respectively, between the two extreme values of the FWHM.

### 2.3. $D_n$ model with two or more channels of infection

Some models [22] include changes in the transmission rate due to various interventions implemented to contain the outbreak. The simple  $D$  model does not allow to do this explicitly, but changes in the spread can be taken into account by considering the total  $D$  or  $D_n$  function as the sum of two or more independent  $D$ -functions with different parameters, which may reveal the existence of several independent sources, or virus channels. An example is shown in Fig. 5, where the two-channel function

$$D'_2 = D'(a, b, c) + D'(a_2, b_2, c_2), \quad (16)$$

has been fitted with six parameters to the Spanish data up to April 13. The fit reveals a second, smaller death peak, which substantially increase the number of deaths per day and the duration of the pandemic. This is equivalent to add a second,

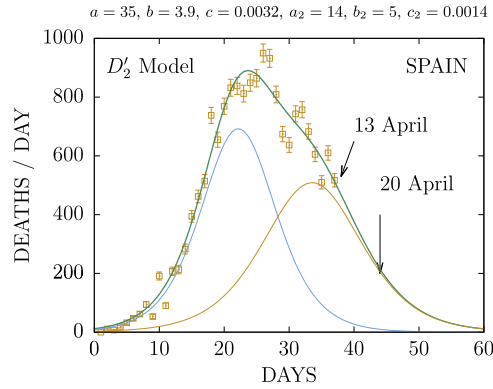


Fig. 5. Predictions of the  $D_2$  model in Spain using a sum of two  $D'$ -functions for data collected up to April 13.

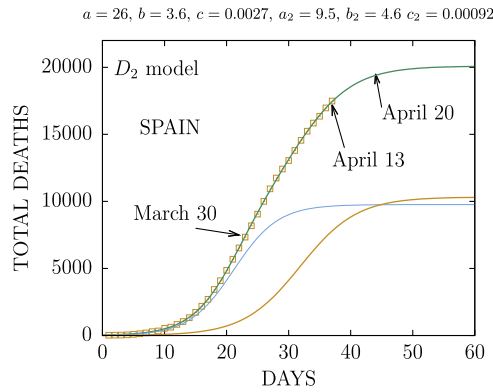


Fig. 6. Predictions of the  $D_2$  model in Spain using a sum of two  $D$ -functions for data collected up to April 13.

independent, source of infection several weeks after the initial pandemic. The second peak may as well represent a second pandemic phase driving the effects of quarantine during the descendant part of the curve.

Additionally, the cumulative  $D$ -function can also be computed with a two-channel function,

$$D_2 = D(a, b, c) + D(a_2, b_2, c_2), \quad (17)$$

which provides, as shown in Fig. 6, a more accurate prediction for the total number of deaths and clearly illustrates the separate effect of both source peaks. It is interesting to note that for large  $t$ ,  $a \approx a_2$ ,  $c \approx c_2$  and  $b_2 \approx 2b$ . In such a case, the total number of deaths expected during the pandemic is given by  $D_2(\infty) = 2a/c$ .

#### 2.4. Estimation of the infected function $I(t)$

The  $D$ -model can also be used to estimate  $I(t)$  using the initial values of  $I_0 = I(0)$  and the total number of susceptible people  $N = S(0)$ . The initial value of  $N$  is unknown, and not necessarily equal to the population of the whole country since the pandemic started in localized areas. Here, we shall assume  $N = 10^6$ , although plausible values of  $N$  can be tens of millions. Note that the no-recovery assumption of the  $D$  model is unrealistic, and this calculation only provides an estimation of the number of individuals that were infected at some time, independently of whether they recovered or not.

From the definition of  $D(t)$  in Eq. (9), the following relations between the several parameters of the model were extracted

$$a = \mu I_0 e^{-\tau/b}, \quad (18)$$

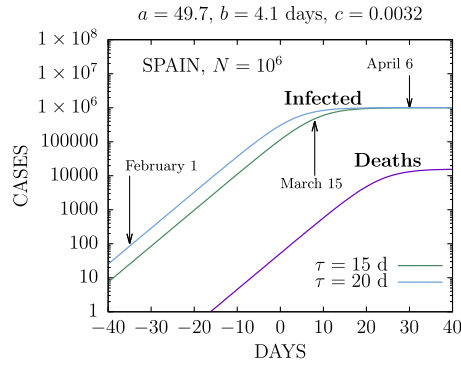
$$c = \frac{I_0}{N} e^{-\tau/b}, \quad (19)$$

$$b = \frac{1}{\lambda N}. \quad (20)$$

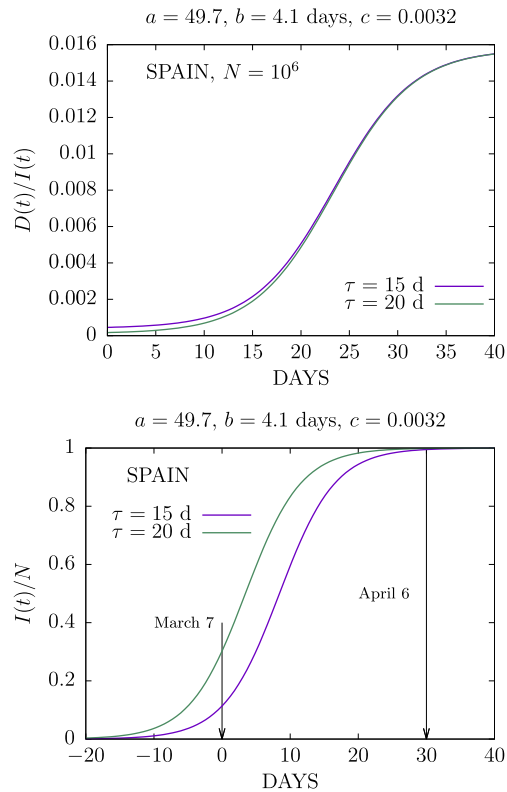
Solving the first two equations for  $\mu$  and  $I_0$  we obtain

$$I_0 = N c e^{\tau/b}, \quad (21)$$

$$\mu = \frac{a}{N c}. \quad (22)$$



**Fig. 7.** Predictions of the D model for the infected function  $I(t)$  in Spain according to data collected up to April 6.



**Fig. 8.** (Top panel) Results for  $D(t)/I(t)$  (deaths over infected) and (bottom panel)  $I(t)/N$  (infected over susceptible) according to the D-model. Results are for data collected in Spain up to April 6, assuming  $N = 10^6$ .

Hence,  $\mu$  can be computed by knowing  $N$ . However, to obtain  $I_0$  one needs to know the death time  $\tau$ . This has been estimated to be about 15–20 days for COVID-19 cases, which can be used to compute two estimates of  $I(t)$ . These are given in Fig. 7 for the case of Spain.

Since there is no recovery in the D model, the total number of infected people is  $I \sim N$  for large  $t$ , i.e.  $N = 10^6$  in our case. In Fig. 7, we have labeled the beginning of the lockdown in Spain (March 15). For  $\tau = 15$  days, most of the susceptible individuals were already infected on that date, and even more for  $\tau = 20$  days, as the pandemic had started almost two months earlier. Most of the individuals got infected, even if a great part of them – approximately 99% – had no symptoms of illness or disease.

Moreover, the top panel of Fig. 8 shows the ratio  $D(t)/I(t)$  (deaths over infected), as given by Eqs. (8) and (10),

$$\frac{D(t)}{I(t)} = \frac{a}{Nc} \frac{1 + ce^{(t+\tau)/b}}{e^{\tau/b} + 1 + ce^{\tau/b}}, \quad (23)$$



which also depends on  $N$  and  $\tau$ . For  $N = 10^6$ , the ratio  $D/I$  increases similarly to the separate functions  $D$  and  $I$  between the initial and final values,

$$\frac{D(0)}{I(0)} = \frac{a}{Nc e^{\tau/b}}, \quad (24)$$

$$\frac{D(\infty)}{I(\infty)} = \frac{a}{Nc}. \quad (25)$$

These results depend on the total susceptible population  $N$ . However, the ratio of infected with respect to susceptibles,  $I/N$ , is independent on  $N$ . This function depends only on  $\tau$  and is shown in the bottom panel of Fig. 8 for  $\tau = 15$  and 20 days, which reveals the rapid spread of the pandemic. Accordingly, between 10% and 30% of the susceptibles were infected in March 7, and one month later (April 6), when the fit was made, all susceptibles had been infected. This does not mean that the full population of the country got infected, since the number  $N$  is unknown and, for instance, excludes individuals in isolated regions, and it may additionally change because of spatial migration, not considered in the model.

### 3. The extended SIR model

D-model predictions can be compared with more realistic results given by the complete SIR model [9,11], which is characterized by Eqs. (1), (2), (3) and (4) with initial conditions  $R(0) = 0, I(0) = I_0, S(0) = N - I_0$ . The SIR system of dynamical equations can be reduced to a non-linear differential equation. First, dividing Eq. (1) by Eq. (3) one obtains,

$$\frac{dS}{dR} = -\frac{\lambda}{\beta} S, \quad (26)$$

which yields the following exponential relation between the susceptible and the removed functions,

$$S = S_0 e^{-\lambda R/\beta}. \quad (27)$$

Moreover, Eq. (4) provides a relation between the infected and the removed functions,

$$I = N - S - R = N - S_0 e^{-\lambda R/\beta} - R, \quad (28)$$

which yields, by inserting into Eq. (3), the final SIR differential equation

$$\frac{dR}{dt} = \beta (N - S_0 e^{-\lambda R/\beta} - R). \quad (29)$$

In order to obtain  $R(t)$  we only need to solve this first-order differential equation with the initial condition  $R(0) = 0$ . Moreover, if we normalize the functions  $S, I$  and  $R$  to 1,

$$S = sN, \quad (30)$$

$$I = iN, \quad (31)$$

$$R = rN, \quad (32)$$

so that  $s + i + r = 1$ , then  $r(t)$  verifies

$$\frac{dr}{dt} = \beta (1 - s_0 e^{-\lambda N r/\beta} - r), \quad (33)$$

which can be solved numerically, or by approximate methods in some cases. In Ref. [9], a solution was found for small values of the exponent  $\lambda N r/\beta$ . For the coronavirus pandemic, however, this number is expected to increase and be close to one at the pandemic end.

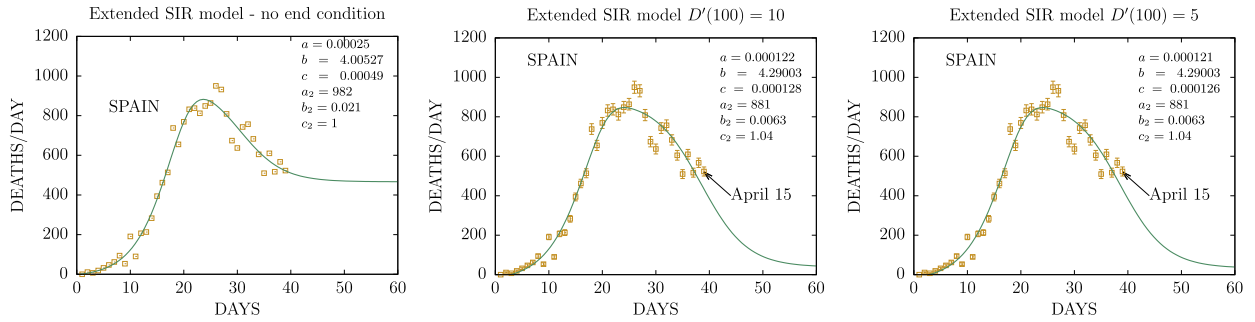
At this point, we propose a modification of the standard SIR model. Instead of solving Eq. (33) numerically and fitting the parameters to data, the solution can be parametrized as

$$r(t) = \frac{a}{c + e^{-t/b}}, \quad (34)$$

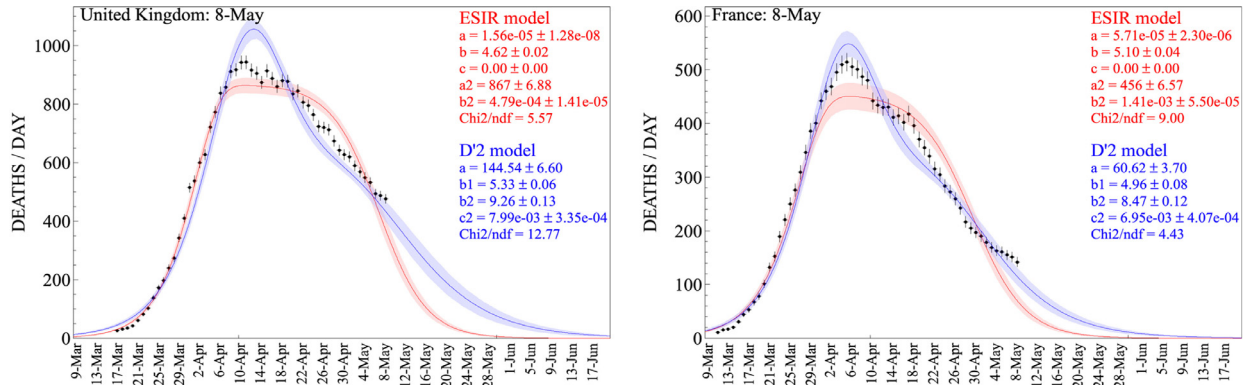
which presents the same functional form as the D-model and, conveniently, provides a faster way to fit the model parameters by avoiding the numerical problem of solving Eq. (33). In fact, numerical solutions of the SIR model present a similar step function for  $R(t)$ . Additionally, one can assume that  $D(t)$  is proportional to  $R(t)$ , and can also be written as

$$\frac{dD}{dt} = a_2 (1 - c_2 e^{-r/b_2} - r(t)), \quad (35)$$

where  $a_2, c_2 = s_0$  and  $b_2 = \beta/(\lambda N)$  are unknown parameters to be fitted to deaths-per-day data, together with the three parameters of the  $r(t)$ -function:  $a, b, c$ .



**Fig. 9.** Fit of the ESIR model to daily deaths in Spain up to April 15 using no boundary condition for the final number of deaths (left panel), and with boundary conditions of  $D'(100) = 10$  (middle panel) and  $D'(100) = 5$  (right panel) deaths/day.



**Fig. 10.** Fit to the data (average of 7 consecutive days up to May 8) of the ESIR and  $D_2$  models in the United Kingdom (left) and France (right).

Fig. 9 shows fits of the ESIR model to daily deaths in Spain during the coronavirus spread. The use of no boundary condition for the number of deaths (left panel) is not an exact solution of the SIR differential equation. A way to solve this problem is to impose the condition  $D'(\infty) = 0$ , as the number of deaths must stop at some time. Numerically, it is enough to choose a small value of  $D'(t)$  for an arbitrary large  $t$ . The middle and right panels of Fig. 9 show different boundary conditions of  $D'(100) = 10$  and  $D'(100) = 5$ , respectively, which yield the same results and the expected behavior for a viral disease spreading and declining. It is also consistently observed (e.g. see middle and right panels of Fig. 9), that at large  $t$ ,  $r(t) \rightarrow \frac{a}{c} \approx 1$ , which essentially means that most of the susceptible population  $N$  recovers, as we previously inferred from the  $D$  model. As shown in Fig. 10, the ESIR model, where  $c_2$  has been adjusted to 1, is characterized by a broad plateau structure which, again, does not consider additional spatial-time effects.

As previously done with the  $D$  model, one can also expand the ESIR model to accommodate its failure to take additional spatial-time effects into account. Similarly, the ESIR2 model is proposed as,

$$\text{ESIR2}(t) = a_2(1 - c_2 e^{-r/b_2} - r(t)), \quad (36)$$

with

$$\begin{aligned} r(t) &= \frac{a}{c + e^{-t/b}} + \frac{a'}{c' + e^{-t/b'}} \\ &= \frac{a}{2a + e^{-t/b}} + \frac{a}{2a + e^{-t/b'}}, \end{aligned} \quad (37)$$

where we have assumed that  $a = a'$  and  $c = 2a$  to accommodate that  $r(\infty) \rightarrow 1$  and  $c_2 = 1$ . Hence, we are left with five free parameters.

Fig. 11 shows the comparison between the ESIR2 and  $D_2$  fits to real data for some European countries where COVID-19 has widely spread: Germany, Italy, France, Spain, United Kingdom, Sweden, Belgium and Netherlands, which indicate a common pattern for the evolution of the COVID-19 pandemic. Death data are taken from Refs. [19,23,24] and consider 7-day average smoothing to correct for anomalies in data collection such as the typical weekend staggering observed in various countries, where weekend data are counted at the beginning of the next week. Real error intervals are extracted from the correlation matrix. As discussed in Section 2.3, the reduced  $D_2$  model has been used with  $a = a_2$  and  $c = c_2$ . Although arising from different assumptions – no recovery ( $D$ -model) and recovery (ESIR model) – both models provide similar patterns following the data trends, with slightly better values of  $\chi^2$  per degree of freedom for the ESIR2 model. It is also interesting to note that the reduced ESIR2 model with five parameters yields similar results to the full ESIR2 model, with eight parameters.

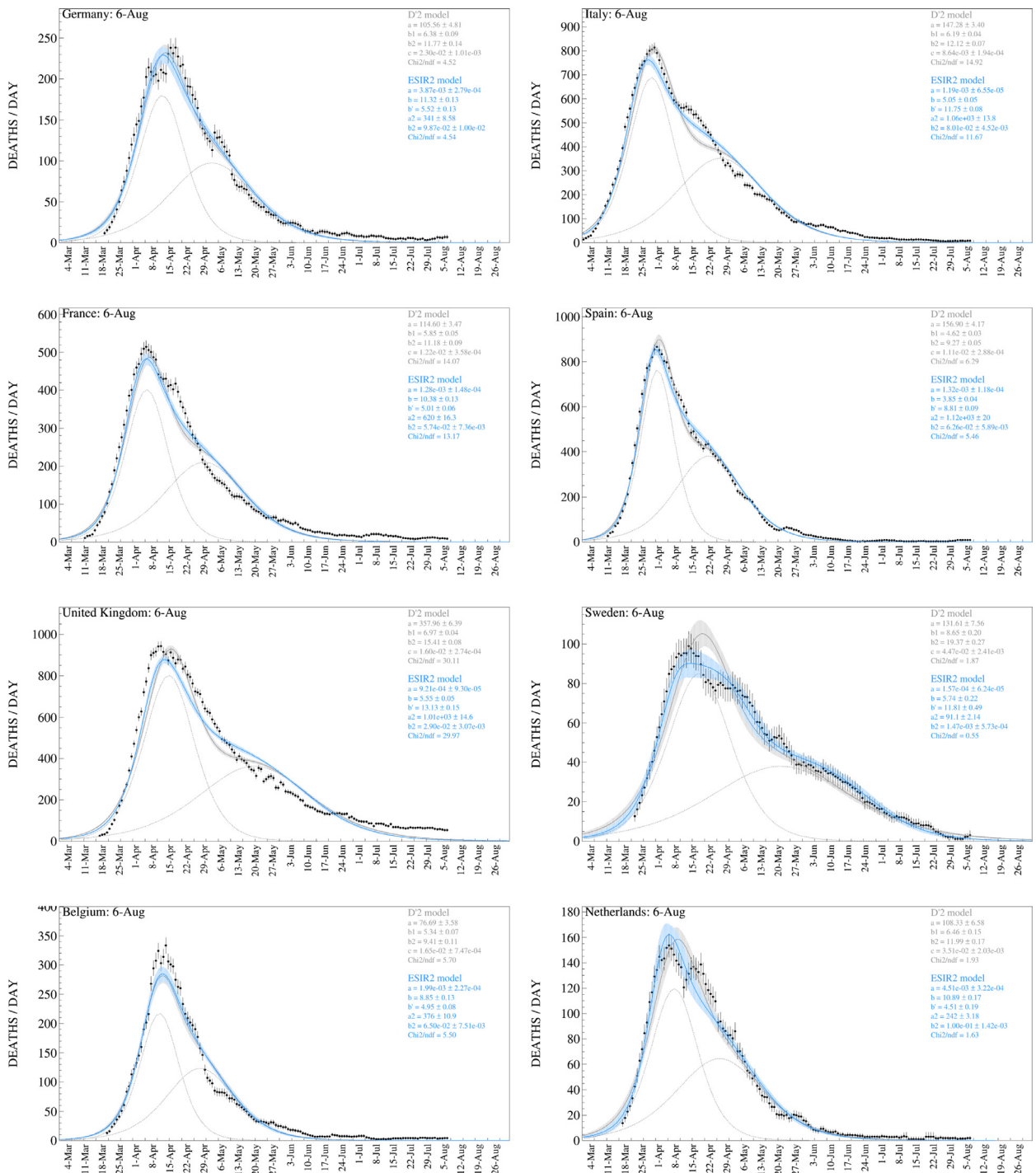


Fig. 11. Reduced ESIR2 and  $D_2$  model fits to deaths-per-day data up to August 9.

As data become available, daily predictions vary for both ESIR2 and the  $D_2$  models. This is because the model parameters are actually statistical averages over space-time of the properties of the complex system. No model is able to predict changes over time of these properties if the physical causes of these changes are not included. The values of the model parameters are only well defined when the disease spread is coming to an end and time changes in the parameters have little influence. Contrarily, Fig. 12 shows clear discrepancies between  $D_2$  and ESIR2 fits to data with larger  $\chi^2/ndf$  values. There are several reasons for these anomalies: (1) a second wave surges as lockdown measures are suddenly released, as clearly shown in

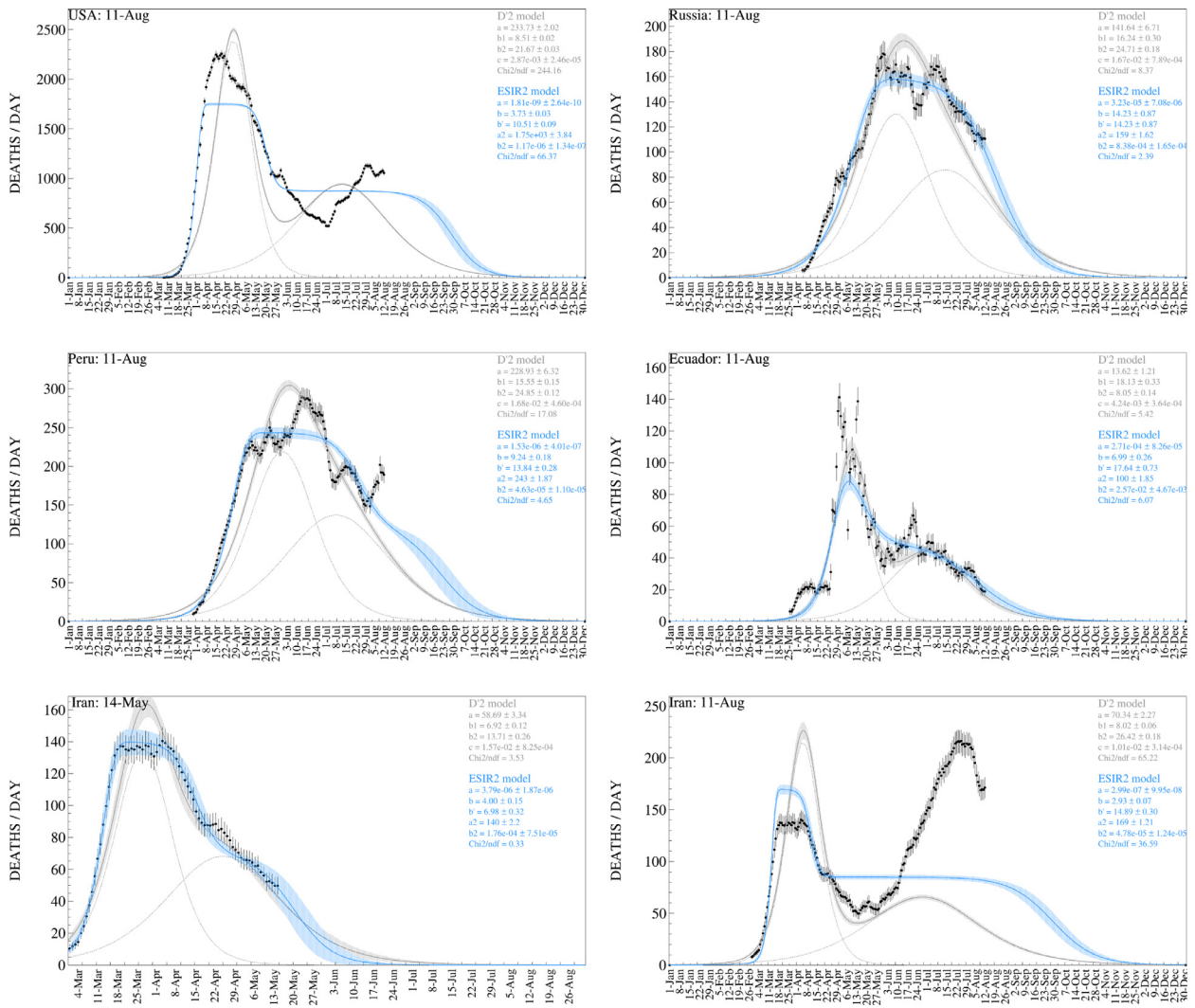


Fig. 12. Discrepancies between the reduced  $D_2$  and ESIR2 models and deaths-per-day data.

the case of Iran, (2) different spatial-time effects as the virus spreads throughout a large country, or, generally, (3) simply because of defective counting (e.g. weekend and backlog effects).

#### 4. Discussion of global results

More sophisticated calculations can be compared with  $D_2$  and ESIR2 predictions. In particular, Monte Carlo (MC) simulations have also been performed in this work for the Spanish case [25], which consist of a lattice of cells that can be in four different states: susceptible, infected, recovered or death. An infected cell can transmit the disease to any other susceptible cell within some random range  $R$ . The transmission mechanism follows principles of nuclear physics for the interaction of a particle with a target. Each infected particle interacts a number  $no$  of times over the interaction region, according to its energy. The number of interactions is proportional to the interaction cross section  $\sigma$  and to the target surface density  $\rho$ . The discrete energy follows a Planck distribution law depending on the ‘temperature’ of the system. For any interaction, an infection probability is applied. Finally, time-dependent recovery and death probabilities are also applied. The resulting virus spread for different sets of parameters can be adjusted from COVID-19 pandemic data. In addition, parameters can be made time dependent in order to investigate, for instance, the effect of an early lockdown or large mass gatherings at the rise of the pandemic.

As shown in Fig. 13, our MC simulations present similar results to the  $D_2$  model, which validates the use of the simple D-model as a first-order approximation. More details on the MC simulation will be presented in a separate manuscript [25]. Interestingly, MC simulations follow the data trend up to May 11 without any changes in the parameters for nearly two

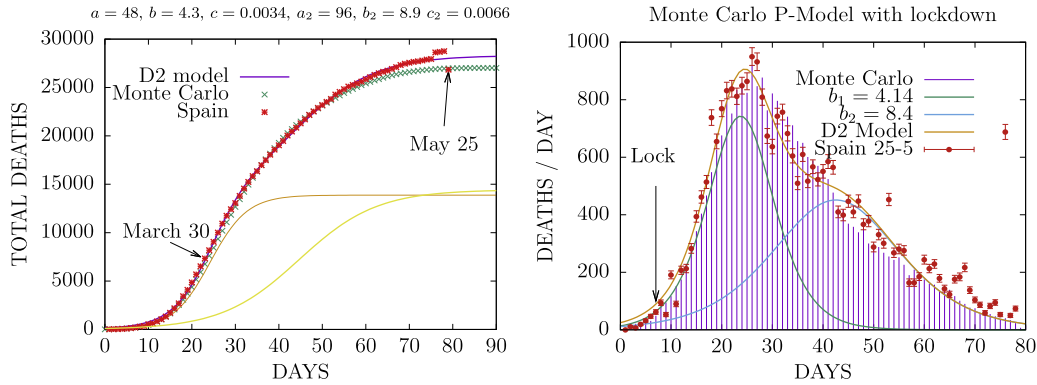
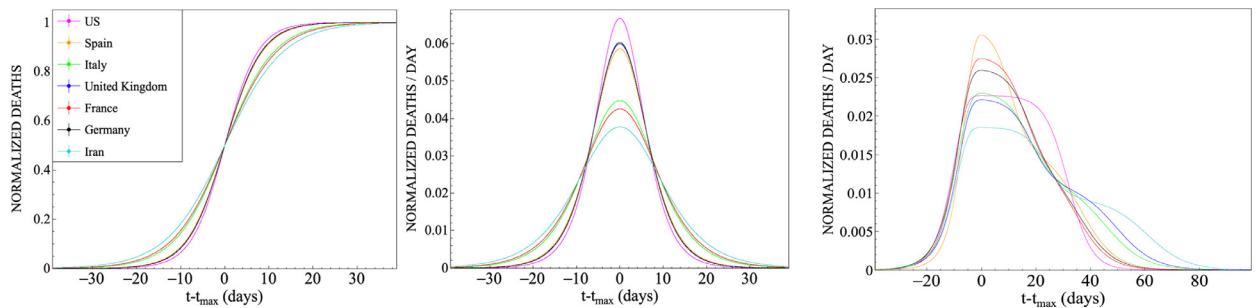


Fig. 13. Predictions of the MC and D models in Spain up to May 11.

Fig. 14. Universality of the normalized  $D$ (left),  $D'$ (middle) and ESIR2 (right) models.

weeks. An app for Android devices, where the Monte Carlo Planck model has been implemented to visualize the simulation is available from Ref. [26].

In order to investigate the universality of the pandemic, it is interesting to compare all countries by plotting the  $D$  model in terms of the variable  $(t - t_0)/b$ , where  $t_0$  is the maximum of the daily curve given by  $t_{\max} = -b \ln(c)$ . By shifting Eq. (10) by  $t_{\max} = -b \ln(c)$  and dividing by  $t_{\max} = a/c$ , the normalized  $D$  function is given by,

$$D_{\text{norm}}(t) = \frac{c e^{(t-t_{\max})/b}}{1 + c e^{(t-t_{\max})/b}}. \quad (38)$$

The left of Fig. 14 shows similar trends for the normalized  $D$  curves of different countries, which suggests a universal behavior of the COVID-19 pandemic. Only Iran seems to slightly deviate from the global trend, which may indicate an early and more effective initial lockdown. A similar approach can be done for the daily data using the  $D'$  and ESIR2 models, as shown in the middle and right panels of Fig. 14, respectively. Although different countries show similar trends, statistical fluctuations in the daily data do not result in a nice universal behavior as compared with  $D_{\text{norm}}$ . However, the  $D'$  and ESIR2 plots show that an effective lockdown is characterized by flatter and broader peaks, best characterized the Iranian case, whereas Spain and Germany present the sharper peaks.

## 5. Final remarks

The global models considered in this work present some differences with respect to other existing models. First, in this work we have tried to keep the models as simple as possible. This allows to use theoretical-inspired analytical expressions or semi-empirical formulae to perform the data analysis. The use of semi-empirical expressions for describing physical phenomena is recurrent in physics. One of the most famous is the semi-empirical mass formula from nuclear physics. Of course the free parameters need to be fitted from known data, but this allowed to obtain predictions for unknown elements.

In our case we were inspired by the well known statistical SIR-kind models slightly modified to obtain analytical expressions that carry the leading time dependence.

We have found that the  $D$  and  $D_2$  models allow a fast and efficient analysis of the pandemics in the initial and advanced stages. Our results show that the time dependence of the pandemic parameters due to the lockdown can be effectively simulated by the sum of two  $D$ -functions with different widths and heights and centered at different times. The distance between the maxima of the two  $D$ -functions should be a measure of the time between the effective pandemic beginning and lockdown.



In the Spanish case this is about 20 days. Taking into account that lockdown started in March 14, this marks the pandemic starting time as about February 22. Had the lockdown started on that date, the deaths would have been highly reduced. The smooth blending between the two peaks provides a transition between the two statistical regimes (or physical phases) with and without lockdown.

The Monte Carlo simulation results are in agreement with our previous analysis with the  $D$  and  $D_2$  models. The Monte Carlo generates events in a population of individuals in a lattice or grid of cells. We simulate the movement of individuals outside of the cells and interactions with the susceptible individuals within a finite range. The random events follow statistical distributions based on the exponential laws of statistical mechanics for a system of interacting particles, driven by macroscopic magnitudes as the temperature, and interaction probabilities between individuals, that can be related to interaction cross sections.

The Monte Carlo simulation spread the virus in space-time, and allows also space-time dependence on the parameters. In this work we have made the simplest assumptions, only allowing for a lockdown effect by reducing the range of the interaction starting on a fixed day. This simple modification allowed to reproduce nicely the Spanish death-per-day curve. The lockdown produces a relatively long broadening of the curve and a slow decay. Similar MC calculations can be performed in several countries to infer the devastating effect of a late lockdown as compared with early lockdown measures. The later is the case of South Africa and other countries, which have not reached the exponential growth.

The Death (D) and extended SIR (ESIR) models are simple enough to provide fast estimations of pandemic evolution by fitting spatial-time average parameters, and present a good first-order approximation to understand secondary effects during the pandemic, such as lockdown and population migrations, which may help to control the disease. Similar models are available [17,27], but challenges in epidemiological modeling remain [28–31]. This is a very complex system, which involves many degrees of freedom and millions of people, and even assuming consistent disease reporting – which is rarely the case – there remains an important open question: Can any model predict the evolution of an epidemic from partial data? Or similarly, Is it possible, at any given time and data, to measure the validity of an epidemic growth curve? We finally hope that we have added new insightful ideas with the Death, the extended SIR and Monte Carlo models, which can now be applied to any country which has followed the initial exponential pandemic growth.

It is important to note that the ESIR and D models predict similar patterns of infected and death cases assuming very different premises: recovery and no recovery, respectively. This, together with the fact that the ESIR model predicts that  $r \rightarrow 1$  for large  $t$ , i.e. that most infected cases eventually recover, leads to the logical conclusion that most people in a fixed population  $N$  become asymptomatic in the D model and eventually recover from COVID-19. One remaining important question is what is  $N$  exactly; is it the whole country, a state or a province, or is it localized to specific areas?

## Acknowledgment

The authors thank useful comments from Emmanuel Clément, Araceli Lopez-Martens, David Jenkins, Ramon Wyss, Azwinndini Muronga, Liam Gaffney and Hans Fynbo. This work was supported by the Spanish Ministerio de Economía y Competitividad and European FEDER funds (grant FIS2017-85053-C2-1-P), Junta de Andalucía (grant FQM-225) and the South African National Research Foundation (NRF) under Grant 93500.

## References

- [1] R.M. Anderson, Discussion: the kermack-mckendrick epidemic threshold theorem, *Bull. Math. Biol.* 53 (1) (1991) 132.
- [2] S. Chauhan, O.P. Misra, J. Dhar, Stability analysis of SIR model with vaccination, *J. Comput. Appl. Math.* 4 (1) (2014) 17–23.
- [3] D.L. Chao, D.T. Dimitrov, Seasonality and the effectiveness of mass vaccination, *Math. Biosci. Eng.* 13 (2) (2016) 249–259.
- [4] H.S. Rodrigues, Application of SIR epidemiological model: new trends, 2016, [arXiv:1611.02565](https://arxiv.org/abs/1611.02565).
- [5] W.H. Hamer, Epidemic disease in England – the evidence of variability and of persistency of type, *Lancet* II (1906) 733–739.
- [6] R. Ross, Report on the Prevention of Malaria in Mauritius, Waterlow and Sons, 1908.
- [7] R. Ross, An application of the theory of probabilities to the study of a priori pathometry. – part I, *Proc. R. Soc. Lond. A* 92 (1916) 204–230.
- [8] R. Ross, H.P. Hudson, An application of the theory of probabilities to the study of a priori pathometry – part III, *Proc. R. Soc. A* 93 (1917) 225–240.
- [9] W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics, *Proc. R. Soc. A* 115 (1927) 700–721.
- [10] D.G. Kendall, Discussion of measles periodicity and community size by M. S. Bartlett *J. Roy. Stat. Soc. A* 120 (1957) 64–76.
- [11] M.S. Bartlett, Measles periodicity and community size, *J. R. Stat. Soc. A* 120 (1) (1957) 48–70.
- [12] M.S. Bartlett, Deterministic and stochastic models for recurrent epidemics, *Berkeley symposium, on Math. Statist. and Prob., Proc. Third Berkeley Symp. on Math. Statist. and Prob. (Univ. of Calif. Press, 4 (1956) 81–109.*
- [13] W.D. Flanders, D.G. Kleinbaum, Basic models for disease occurrence in epidemiology, *Int. J. Epidemiol.* 24 (1) (1995) 1–7.
- [14] J.E. Amaro, The d model for deaths by covid-19, 2020, [arXiv:2003.13747](https://arxiv.org/abs/2003.13747).
- [15] D.S. Hui, E. Azhar, T.A. Madani, The continuing 2019-ncov epidemic threat of novel coronaviruses to global health the latest 2019 novel coronavirus outbreak in wuhan, china, *Int. J. Infect. Dis.* 91 (2020) 264–266.
- [16] WHO, coronavirus disease 2019 (COVID-19) situation report – 83. 12 april, 2020, <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>.
- [17] L. Squillante, I.F. Mello, A.C. Seridonio, M. de Souza, Attacking the COVID-19 with the ising-model and the fermi-dirac distribution function, 2020, [arXiv:2003.11860](https://arxiv.org/abs/2003.11860).
- [18] H. Weiss, The SIR model and the foundations of public health, *MATerials MATematics* no. (3) (2013).
- [19] <https://www.worldometers.info/coronavirus/>.
- [20] N.A. Kudryashov, *Appl. Math. Model.* 39 (2015) 5733.
- [21] P.R. Hunter, F.J. Colón-González, J. Brainard, S. Rushton, Impact of non-pharmaceutical interventions against COVID-19 in Europe: a quasi-experimental study, *MedRxiv* (2020), doi:[10.1101/2020.05.01.20088260](https://doi.org/10.1101/2020.05.01.20088260).
- [22] J. Dehning, Inferring covid-19 spreading rates and potential change points for case number forecasts, 2020, [arxiv:2004.01105](https://arxiv.org/abs/2004.01105).
- [23] <https://covid19.isciii.es/>.

- [24] <https://dashboard.covid19.data.gouv.fr/>.
- [25] J.E. Amaro, J.N. Orce, in preparation (2020).
- [26] <https://www.ugr.es/~amaro/coronavirus/>.
- [27] <https://covid19.healthdata.org/>.
- [28] J. Lloyd-Smith, D. Mollison, J. Metcalf, P. Klepac, H. Heesterbeek (Eds.), Special issue on challenges in modelling infectious disease dynamics, vol. 10, 1–108, 2015. *Epidemics*.
- [29] H. Heesterbeek, Modeling infectious disease dynamics in the complex landscape of global health, *Science* 347 (6227) (2015) aaa4339.
- [30] F. Brauer, Mathematical epidemiology: past, present, and future, *Infect. Dis. Model.* 2 (2017) 113–127.
- [31] S.G. Krantz, P. Polyakov, A.S.R.S. Rao, True epidemic growth construction through harmonic analysis, *J. Theor. Biol.* 494 (2020) 110243.