Covid-19 predictions using a Gauss model, based on data from April 2

Janik Schütter,[1] Reinhard Schlickeiser*,[2,3] Frank Schlickeiser,[2] and Martin Kröger*[4]

[1]*Department of Chemistry and Applied Biosciences, ETH Zurich, 8093 Zurich, Switzerland*
[2]*Institut für Theoretische Physik, Lehrstuhl IV: Weltraum- und Astrophysik, Ruhr-Universität Bochum, D-44780 Bochum, Germany*
[3]*Institut für Theoretische Physik und Astrophysik, Christian-Albrechts-Universität zu Kiel, D-24118 Kiel, Germany*†
[4]*Polymer Physics, Department of Materials, ETH Zurich, Switzerland*‡

(Dated: April 5, 2020)

We propose a Gauss model (GM), a map from time to the bell-shaped Gauss function to model the casualties per day and country, as a quick and simple model to make predictions on the coronavirus epidemic. Justified by the sigmoidal nature of a pandemic, i.e. initial exponential spread to eventual saturation, we apply the GM to the first corona pandemic wave using data from 25 countries, for which a sufficient amount of not yet fully developed data exists, as of April 2, 2020, and study the model's predictions. We find that logarithmic daily fatalities caused by Covid-19 are well described by a quadratic function in time. By fitting the data to second order polynomials from a statistical χ^2-fit with 95% confidence, we are able to obtain the characteristic parameters of the GM, i.e. a width, peak height and time of peak, for each country separately. We provide evidence that this supposedly oversimplifying model might still have predictive power and use it to forecast the further course of the fatalities caused by Covid-19 per country, including peak number of deaths per day, date of peak, and duration within most deaths occur. While our main goal is to present the general idea of the simple modeling process using GMs, we also describe possible estimates for the number of required respiratory machines and the duration left until the number of infected will be significantly reduced.

I. Introduction

Nowadays, many models to predict the spreading of infectious diseases like Covid-19 are available, for example the actively discussed susceptible-infected-removed (SIR) model. State-of-the-art modeling efforts had been summarized very recently elsewhere [3]. Many of these models are either toy models that cannot make reliable predictions or they are so complex, by taking into account a wide range of factors, that simple predictions are not possible. In times of the coronavirus epidemic, *fast* predictions on the course of the coronavirus disease are crucial for policy makers to optimize their managing of the disease wave. To feed into the current debate on infectious disease models, we would like to propose the GM as a simple, but effective description of fatalities caused by Covid-19 over time, similar to a recent study [8] that focused on the analysis of doubling times for Germany reported by a newspaper. In contrast to this previous work, we chose to use the reported daily *death* rates [5] as monitored input data, because in nearly all countries these are better documented than the monitored daily infections and have a significant lower number of unreported cases. We also do not rely on doubling times as their determination is model-dependent.

Predictions such as the maximum number of fatalities per day or the date of the peak number of newly seriously sick persons per day (SSPs) are valuable data for governments around the world, especially those facing the beginning of an exponential increase of casualties, and we hope to serve the people in charge with the here presented simple and swift GM. We here study the predictions for these quantities of interest based on a GM for all those countries for which sufficient – to be made precise below – amounts of data already existed as of April 2nd. We only present the conceptual idea of such a process in form of a sort of recipe and hope to spark motivation for people to quickly recreate GM time evolutions, determine the three model parameters for their country, city, environment, as will be explained in detail, to be able to make predictions for relevant parameters such as the fraction of respiratory machines in use vs. available.

While the GM may appear too simple to be predictive, it does fit the past data well, including the entire first epidemic wave in China, and hence must have at lease some predictive power for data to come. Please note, though, that we are no epidemiologists and have no prior expertise in the modelling of diseases. Also, with the here presented GM we dare, by no means, to present a model capable of similar mechanistic and causal richness compared with existing infectious disease models. Our only addition is to note and use for predictions the macroscopic Gaussian nature of the time evolution of daily fatalities that is universal among all countries and, in fact, almost necessarily so given the sigmoidal nature of total fatalities – a point we will explain in detail in the discussion.

II. Results

A. Gauss model (GM)

We would like to model the time-dependent daily change of infections and daily change of deaths with their own, a priori independent, time-dependent Gaussian functions denoted by $i(t)$ and $d(t)$ in the following. Each Gaussian is a bell-shaped curve, the black line in Fig. 1(a), characterized by three independent parameters: a width, a maximum height and a time at which the Gaussian curve attains this maximum height.

It must be emphasized that we model the *daily* change of deaths, in contrast to the *cumulative* number of deaths, more frequently available in public, since the change of deaths allow for a more stable fit around its maximum which is the time of interest for predicted quantities, cf. the discussion III. The cumulative deaths are the sum of all previous daily deaths up to today, while the number of daily deaths in turn is the difference of two consecutive days in cumulative deaths. In Fig. 1(a), the red plot illustrates the cumulative number of deaths as a function of time for the respective daily number of deaths in the same panel. To appreciate the meaning of the three parameters, the GM curves for varying parameters are displayed in the first row of Fig. 1(b), while the corresponding integrated 'total' versions are shown in the 2nd row.

*Corresponding author
†Electronic address: rsch@tp4.ruhr-uni-bochum.de
‡Electronic address: mk@mat.ethz.ch

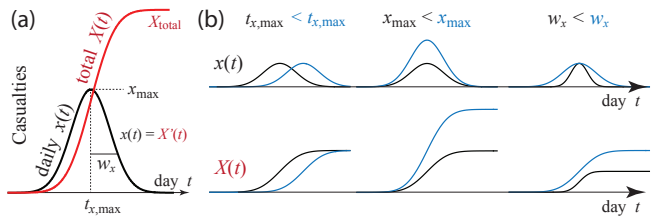


Fig. 1: **(a)** GM for time evolution of a daily quantity $x(t)$ (black) and the corresponding total quantity $X(t)$ (red), which is the cumulative sum of $x(t)$ until time t. **(b)** Consequences of varying the three parameters describing the GM: width w_x, maximum height $x_{\max}$ and time of maximum height $t_{\max}$ for both the daily (top) and total (bottom) rates. In this work x stands for either deaths ($x = d$) or confirmed infections ($x = i$).

B. Logarithmic daily fatalities are quadratic

We went on to fit the time evolution of daily fatalities to the GM in time, one for each country. To do so, we fitted a polynomial of second order to the logarithmic number of daily fatalities of 14 countries as a function of time using a χ^2-fit. The resulting quadratic fit is plotted in Fig. 2. For the remaining 11 countries, similar fits could only be performed for the daily number of infections.

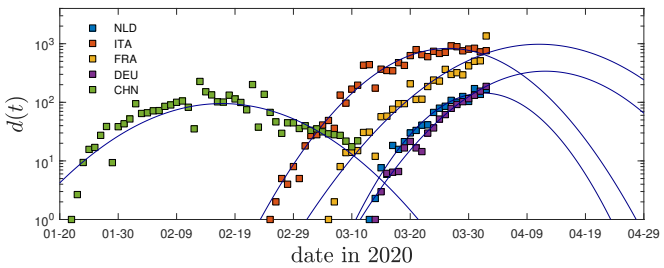


Fig. 2: Logarithmic reported number of daily fatalities (squares) and the quadratic fits of number of daily fatalities (lines) over time for some countries. The plots demonstrate the quadratic nature of the logarithmic fatalities per day. Raw data taken from Ref. [5].

We recall that we prefer to base any quantitative conclusions only on the number of deaths, and not on the number of infections per day. Deaths are better documented than monitored infections in nearly all countries. A death caused by Covid-19 is easier to count than an infection, which might as well cause none to moderate symptoms and hence might remain uncounted. Statistically, a constant fraction of infected die from Covid-19 at a later time after being registered as infected [3, 10]. Thus, infection and death curves are equivalent descriptions of time evolution of Covid-19, and the coefficients characterizing their shape can be expected to be closely related. To demonstrate that both, infections and deaths, follow the GM, we analyzed and show results for both measures.

C. The fitted parameters

Using the fitted polynomial coefficients we can compute the three parameters of the GM, i.e. maximum height, time of maximum height and curve width, for each country. For mathematical details, please refer to the appendix.

To demonstrate the universal Gaussian nature of the daily fatalities over time, we display them in Fig. 3(a,b), normalized so that all curves have unit width, maximum and time of maximum. The same plots for the cumulative fatalities over time are shown in Fig. 3(c,d). Daily infections, daily fatalities, cumulative infections and cumulative fatalities, all fit neatly onto the unit GM curve, plotted in gray in the back for reference. China, which is the only country to provide data from its first pandemic wave for times greater than 0.6 (in normalized units), fits to the GM well over the entire significant course of infections and fatalities. This sparks the hope that the used GM will have predictive power for the remaining countries also after the maximum. The fits already provide sufficient evidence that the part prior to the maximum is captured well by the GM.

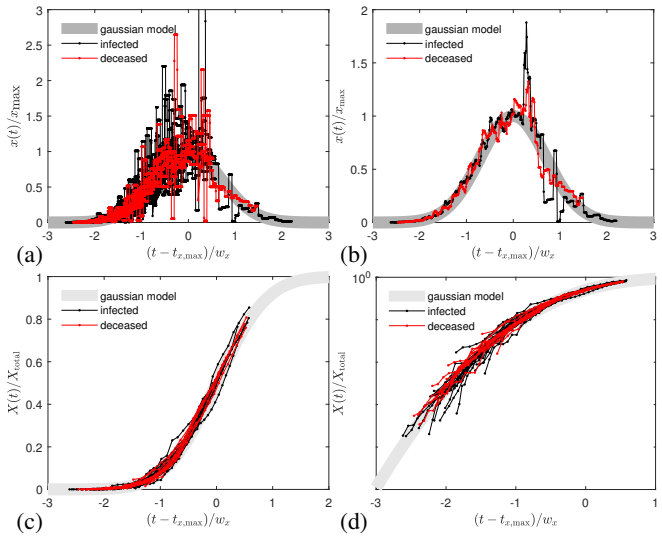


Fig. 3: Shifted and rescaled data for 25 countries. **(a)** number of daily infections and fatalities. **(b)** obtained from (a) by averaging over entries at same $(t - t_{d,\max})/w_d$ (bin size 0.01). **(c)** Number of cumulative infections and fatalities over time in normal scale, and **(d)** in logarithmic scale to appreciate different regimes with better resolution. **(a)-(d)** Thick gray is the theory expression (1) for daily, and (A7) for cumulative casualties. Data beyond $(t - t_{d,\max})/w_d = 0.6$ is from China alone. Raw data taken from Ref. [5].

The resulting GM parameters are listed and plotted in Fig. 4. For most countries the GM width is within 10 and 15 days, roughly half of all countries have passed their peak of daily fatalities already and the peak is roughly below 20 fatalities per day and per million people.

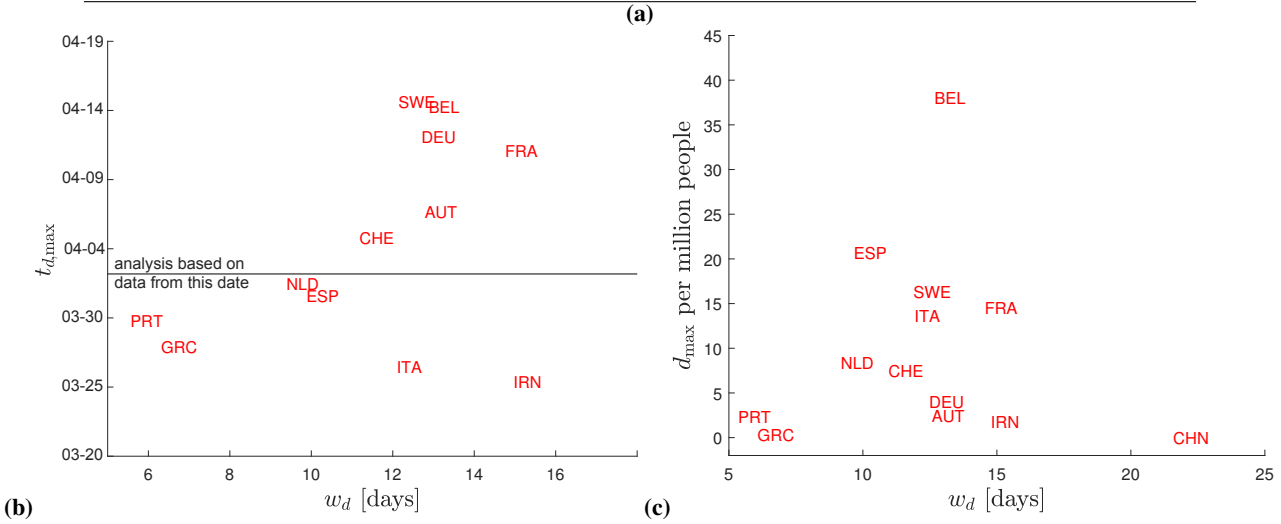
D. Additional predictions

Using the GM, one can obtain predictions for the further course of the Covid-19 pandemic, buy it is equally possible to make speculations about the past. Apart from the width, height and time of maximum directly contained in the GM parameters, predictions for the periods of time T_η relevant for the planning of protective measures are already mentioned in Fig. 4(a). We next present two possible applications out of several others (cf. methods section): cumulative fatalities as a function of time and the maximum required number of respiratory equipment as well as its time point. We hope that the reader uses the following examples as incentive and inspiration to produce their own predictions based on the GM.

First, the time evolution of the number of *cumulative* fatalities, plotted in Fig. 5, can be obtained by summing *daily* number of deaths, predicted by our model. In this figure, we rescaled all curves back to normal times so that the future course of cumulative deaths can be easily read-off. This plot suggests for Italy and Spain to plateau first, while France

α_3 code	Country	w_d	$t_{d,\max}$	$d_{\max}$	D_{total}	D_{total}/MP	$T_{1\%}$	$T_{1‰}$
AUT	Austria	13.2 ± 1.5	04-07 $\pm$ 21	21.6 ± 4.7	500 ± 170	60 ± 20	04-25 $\pm$ 44	05-01 $\pm$ 45
BEL	Belgium	13.3 ± 0.6	04-14 $\pm$ 18	430 ± 26	10200 ± 1100	890 ± 100	05-03 $\pm$ 19	05-09 $\pm$ 19
CHE	Switzerland	11.6 ± 0.5	04-05 $\pm$ 17	63 ± 12	1300 ± 300	156 ± 37	04-20 $\pm$ 18	04-25 $\pm$ 18
CHN*)	China	15.7 ± 0.3	02-17 $\pm$ 3	95 ± 3	2600 ± 100	1.9 ± 0.1	03-11 $\pm$ 4	03-19 $\pm$ 4
DEU	Germany	13.1 ± 0.4	04-12 $\pm$ 10	340 ± 6	7900 ± 400	95 ± 5	04-30 $\pm$ 11	05-07 $\pm$ 11
ESP	Spain	10.3 ± 0.2	04-01 $\pm$ 6	960 ± 70	17500 ± 1600	380 ± 35	04-13 $\pm$ 7	04-18 $\pm$ 7
FRA	France	15.2 ± 0.3	04-11 $\pm$ 8	980 ± 320	26000 ± 9000	390 ± 140	05-04 $\pm$ 9	05-11 $\pm$ 9
GRC	Greece	7.0 ± 0.2	03-27 $\pm$ 9	3.8 ± 1.3	47 ± 17	4.4 ± 1.6	04-01 $\pm$ 7	04-05 $\pm$ 7
IDN	Indonesia	11.9 ± 0.8	03-19 $\pm$ 22	5.3 ± 4.7	111 ± 91	0.4 ± 0.4	—	—
IRN	Iran	15.3 ± 0.1	03-25 $\pm$ 2	150 ± 14	4100 ± 400	51 ± 5	04-17 $\pm$ 2	04-23 $\pm$ 2
ITA	Italy	12.4 ± 0.1	03-27 $\pm$ 1.8	832 ± 60	18300 ± 1400	300 ± 23	04-12 $\pm$ 2	04-18 $\pm$ 2
NLD	Netherlands	9.8 ± 0.1	04-02 $\pm$ 4	144 ± 23	2500 ± 400	147 ± 26	04-13 $\pm$ 5	04-17 $\pm$ 5
PRT	Portugal	6.0 ± 0.1	03-29 $\pm$ 4	24 ± 4	260 ± 40	25 ± 4	04-02 $\pm$ 4	04-05 $\pm$ 4
SWE	Sweden	12.6 ± 1.2	04-15 $\pm$ 35	162 ± 12	3600 ± 600	370 ± 60	05-01 $\pm$ 35	05-08 $\pm$ 40

(a)



(b) **(c)**

Fig. 4: Fitted GM parameters w_d, $t_{d,\max}$, and $d_{\max}$ for those countries, for which sufficient data about fatalities is currently available. Error bars reported for 95% confidence intervals. **(a)** Table listing the three fitted GM parameters, followed by estimated cumulative number of fatalities D_{total}, the same quantity per million people (MP), and the projected dates T_η (A10) by which the number of daily *infected* people had reduced to the level of η of its maximum value. Number of inhabitants according to OECD [1]. **(b)** Plot of the fitted parameters. The horizontal lines marks the day of this study, April 2nd. China is missing on this plot as its $t_{d,\max}$ occurred on about Feb 17 according to our calculation. **(c)** Peak daily number of fatalities per 1 million inhabitants, i.e., $d_{\max}$ divided by the number of inhabitants times 10^6. *) For China we considered only the data during the first pandemic wave, i.e., until a minimum in daily fatalities was clearly reached on March 12.

will have to face increasing number of fatalities considerably longer. It also predicts the cumulative number of fatalities per million people over the entire course of the Covid-19 disease to be highest for France, Spain and Italy.

Next, we estimate the number of required respiratory machines per date for the Covid-19 epidemic. We start by assuming the number of respiratory machines per day to be equal to the cumulative number of *active* seriously sick persons on the given date, where active means not yet recovered by that date. Each new seriously sick person per day (SSP) requires a respiratory machine for some days or even weeks before passing away or recovering from Covid-19. According to other works [10], people to have died from Covid-19 occupied respiratory equipment for an average of 7 days prior to their death, but respiratory equipment may be in use for up to about one month for cases that later recovered. Thus, we may roughly estimate the number of active SSPs per date as the sum of people that became seriously sick within the past 10 days. Please note, however, that we try to only conceptually link the GM to useful quantities, we leave a thorough search of exact numbers to the reader.

As a final step, we would like to relate the SSPs to deaths. Assume that each SSP dies with a constant probability γ after some days, i.e. γ times SSPS(t) gives the daily fatalities some days ahead. Taking again numbers from Ref. [10], we could use that each deceased patient had used a respiratory machine for an average of 7 days prior to death and thus estimate the daily number of SSPs at a given date by the number of daily fatalities 7 days in the future divided by probability γ.

The result of the above estimate reveals that the required number of respiratory equipment itself is a Gaussian curve, roughly centered around the same date as the daily fatality curve, and its peak value is proportional to a multiplicative factor that depends on the width of the Gaussian and ranges between 0.5 and 0.9, the total number of fatalities D_{total}, and the probability of passing away as SSP γ.

III. Discussion

The here presented GM allows for simple predictions of future course of the Covid-19 disease. Using this model and our recipe to extract its parameters, interested readers are in the position to obtain estimates for the shape of the Gaussian curve for their country, state, community, and use this model to compute more quantities of interest, such as our sketch of how to estimate the maximum number of required respiratory machines and the date of this maximum demand. Knowing the

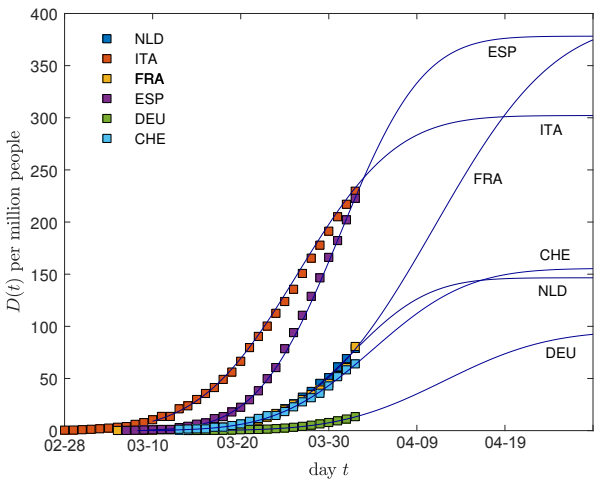


Fig. 5: Measured cumulative number of deaths per one million inhabitants (symbols), compared with the GM predictions (A7) for selected fits of Fig. 4. The cumulative number of fatalities determining the height of the future plateau is given by the product of width w_d and $d_{\max}$, times $\sqrt{\pi} \approx 1.77$.

time of maximum rush days of newly SSPs to their hospitals, the maximum number of newly SSPs and width could help the government and medical agencies in these countries to optimize the managing of the disease wave by appropriate drastic actions for limited time e.g. as mobilizing the army for help in the hospitals. Moreover, fortunately, as our study here demonstrates, the time of peak of the disease wave differs among countries. Knowledge of these peak times times and their durations allows other countries to help those who undergo the peak of the wave at a significantly later time, with breathing apparati and trained medical personal for a brief predictable time.

The are various other 'microscopic' models that can lead to a GM dynamics of fatalities or infections. One of more prominent ones recently appeared in the Washington Post [9]. Stevens investigated what happens when simulitis spreads in a town, if everyone in the town starts at a random position, moving at a random angle, infecting others upon collision, and recovering after a certain time. The simulated number of infected people rises rapidly as the disease spreads and tapers off as people recover. We recreated the simulations and found evidence for the applicability of the GM under many circumstances. These results are not reported here, but support our central assumption. From another recent work using a holistic agent-based model [2], where the agents adapt their behavior through artificial intelligence as part of the solution, there seems also evidence from the numerical results presented, that the number of newly infected may be well captured by a Gaussian function.

The question remains, though, how the GM could be justified? Intuitively, we know that the cumulative number of casualties for a wave of any pandemic must start from a constant (often 0), then increase exponentially and eventually saturate at a higher constant level. Functions that capture such a behavior, i.e. a smooth change from a lower constant to a higher constant over a finite duration, are called sigmoidal. The derivative of sigmoidal functions have a bell-shaped form, similar to a Gaussian function, but in general can be asymmetric. We here model the daily fatalities, formally the derivative of the cumulative fatalities. Since we expect the cumulative fatalities to be sigmoidal, from common-sense reasoning as argued above, this fixes the derivative, the daily fatalities, to a bell-shaped form.

Even though all pandemics thus give rise to bell-shaped change in casualties, the curve's parameters might differ, influenced mostly by policy, health system and culture. The predictive power of our model rests on the assumption that these influences are encoded already into the early data of casualties, combined with the assumption that the principal shape of all pandemics is fixed.

Why do we choose a symmetric bell-shaped form, the Gaussian function? Three aspects: fist a symmetric function is the simplest model among all bell-shaped functions and thus suffices to convey the idea of such models. Second, it works, in the sense discussed here. And third, the times of greatest interest to policy makers are until the bell-curve's peak since once passed the health system should be able to cope.

Sigmoidal models for predictions on the course of a pandemic are not new [6, 7]. For example, Fu et al. [6] used a logistic function as another instance of a sigmoidal function. In our experiments we found such fits to be generally sensitive to initial conditions and to often require a large number of parameters. We therefore chose to fit the logarithm of daily change of cases instead of the cumulative cases. The logarithm of daily change of cases weights more strongly values close to the functions maximum and disregards other values. This leads to a more stable model of the daily change of cases around its maximum, the turning point of the cumulative cases and the time of interest since most relevant predictions such as peak of the pandemic, time point and width of peak are focused around it.

Compared to our previous study [8], that predicted the peak of the first pandemic wave in Germany to be April 11th, 2020 $^{+5.4}_{-3.4}$ and with a delay of about 7 days the maximum demand on breathing machines in hospitals occurs in Germany on April 18th, 2020 $^{+5.4}_{-3.4}$ days. These predictions are in accordance with the prediction from the analysis in this paper, see the data for Germany in Fig. 4. However, in contrast to our methodology, the authors rely on doubling times for their predictions. These doubling times differ depending on the way they are calculated, from daily casualties or cumulative casualties, and is almost all cases require preprecessing such as smoothing (methods section).

IV. Conclusion

In this document we have provided some evidence that a GM may be used to capture the time evolution of the daily fatalities and infections per country. Fitted models describe past data well, including data from China. How is this GM useful? Our hope is to guide others into using such model. The model is so simple that it can be reproduced and applied without detailed knowledge of epidemiology, statistics or programming languages. There are many countries not yet drastically affected by Covid-19, which will likely change for many in the coming weeks, and the GM could for example be used to apply it to such countries as soon as sufficient data is available. Besides that we hope to make the public aware of the Gaussian or sigmoidal nature emerging from Covid-19 infections, similar to the numerous discussions of exponential growth in recent times. No pandemic is ever exponential, in the long run it is sigmoidal, and thus makes for a good discussion.

On one hand we are afraid our predictions will become reality, on the other they are more optimistic than all (few) predic-

tions we came across so far. Confronting these predictions and the method with reality will help to either establish or rule out the presented approach within a very short time. It is the simplicity of the model and its missing freedom which will allows us to quickly decide on its usefulness for future applications.

We conclude with a word of caution. We are certainly no experts in this field and a GM is simply a description of a smooth time evolution of infections. We leave it to the reader to treat the here presented observations and claims with enough care.

A. Methods

In this section we make the concepts used intuitively in the main text rigorous by introducing the necessary mathematical language.

Gauss model — Denote the number of daily fatalities as a function of time by $d(t)$ and the cumulative number of fatalities by $D(t)$. We modelled the time evolution of fatalities using the GM, i.e. a Gaussian function of time,

$$d(t) = d_{\text{max}} \exp\left\{-\frac{(t-t_{\text{d,max}})^2}{w_d^2}\right\}, \tag{A1}$$

where w_d denotes the width of the Gaussian, d_{max} denotes the maximum value of fatalities and $t_{\text{d,max}}$ the time point at which this maximum is attained. The identical model and notation applies to the number of daily infections $i(t)$, cumulative infections $I(t)$ and parameters i_{max}, $t_{\text{i,max}}$, w_i.

We used publicly available data of *monitored* cumulative death rates $D_m(t)$, where the subscript m is used to distinguish data from our model, to derive the daily death rates by taking the first time derivative

$$d_m(t) = \frac{\mathrm{d}D_m(t)}{\mathrm{d}t} \tag{A2}$$

and calculated its natural logarithm $\ln d_m(t)$. The GM dynamics (A1) implies

$$\ln d_m(t) = \ln d_{\text{max}} - \left(\frac{t-t_{\text{d,max}}}{w_d}\right)^2 = c_0 + c_1 t + c_2 t^2, \tag{A3}$$

which is a polynomial function of degree 2 with coefficients

$$c_0 = \ln d_{\text{max}} - \frac{t_{\text{d,max}}^2}{w_d^2}, \quad c_1 = \frac{2t_{\text{d,max}}}{w_d^2}, \quad c_2 = \frac{1}{w_d^2}. \tag{A4}$$

The relevant parameters determining the number of deaths per day, the width of the distribution, as well as the position of the peak are then given by

$$d_{\text{max}} = e^{c_0 + c_1^2/(4c_2)}, \quad t_{\text{d,max}} = \frac{c_1}{2c_2}, \quad w_d = \frac{1}{\sqrt{c_2}}. \tag{A5}$$

Fitting & errors — Using a second order polynomial fit to the data we obtained the coefficients c_0, c_1, c_2 as well as their confidence intervals. For this, the Matlab^{R} function `[P,S,M]=polyfit(t,log(I_m),2)` on the natural logarithm of the monitored death rates $\ln I_m(t)$ yields the coefficients P=[c_2,c_1,c_0] of the fit as well as information about the confidence intervals. We made use of the function `polyparci` that uses only core Matlab^{R} functions and does not require the Statistics Toolbox. It uses the procedures outlined in the `polyfit` documentation to calculate the covariance matrix, and uses functions `betainc` and `fzero` to calculate the cumulative t-distribution and the inverse t-distribution for a given probability and degrees-of-freedom. Within the limited amount of time we had to prepare this document, we were unable to compare error estimates from different approaches.

Deaths vs. infections — We have applied the same procedure to the measured number of infected people, $i_m(t)$, giving rise to another set of parameters i_{max}, $t_{\text{i,max}}$, and w_i. We found that the GM widths for infections w_i and fatalities w_d are similar in magnitude, within errors, and that $\tau_{d,\text{max}}$ and $\tau_{i,\text{max}}$ differ by a number of days $\tau \approx 10$ [4], that can be considered constant for practical purposes. Our analysis confirms this estimate. It is also useful to introduce the fraction of fatalities among the truly infected (not the reported infected) $f = D_{\text{total}}/I_{\text{total}}$, as this fraction can be expected to vary within limited bounds. We thus write

$$i_{\text{max}} = \frac{d_{\text{max}}}{f}, \qquad t_{i,\text{max}} = t_{d,\text{max}} - \tau, \qquad w_i \approx w_d \tag{A6}$$

This reduces the number of parameters for a combined study of daily deaths and infections to four, as f cannot be considered constant, or further down to three, employing $f \approx 5\times10^{-3}$ suggested by Fig. 1 of Ref. [4]. We did not make use of these relationships and numbers anywhere in this work, but they can still be used to estimate quantities mentioned below. While this study mostly focuses on the number of fatalities, we had also included data from 11 countries for the reported number of infections in some of the previous figures that provide evidence for the applicability of the GM. Table I lists the corresponding parameters.

Data used — Only countries which as of April 2nd, have reported more than 20 infected or 7 deceased people for more than 10 days. Also, outliers that are better described by a multimodal extension of the GM have been omitted (including the United States) with the exception of China, for which there was a clear end of the first wave on about March 12. This resulted in the 25 counties used here. Using the identical approach, many more countries will be available for analysis within the next few days.

Cumulative fatalities — The accumulated number of fatalities at time t, which we refer to as cumulative number of fatalities, is the integral of the daily fatalities (A1)

$$D(t) = \int_{-\infty}^{t} d(t')\,\mathrm{d}t' = \frac{D_{\text{total}}}{2}\left[1 + \text{erf}\left(\frac{t-t_{d,\text{max}}}{w_d}\right)\right], \tag{A7}$$

where $D_{\text{total}} = d_{\text{max}} w_d \sqrt{\pi}$ is the projected total number of fatalities at $t \to \infty$ and erf is the error function. Using (A7) the time t_0 by which a first patient died from the virus is immediately estimated via $D(t_0) = 1$. Similarly via $I(t_0) = 1$ for the first infected person, so-called patient 0, if one takes into account a time shift τ and ratio f between d_{max} and i_{max}, cf. (A6), and one ignores the fact that the gaussian is likely to break down in this limit. The explicit expression is $t_0 = t_{i,\text{max}} - w_i\,\text{erf}^{-1}(1 - 2/I_{\text{total}})$ for the time of first appearance of Covid-19, and this time is specific for each country. Here, erf^{-1} is the inverse error function. For values close to unity it is well approximated by $\text{erf}^{-1}(1-\epsilon) \approx \sqrt{-\ln(1-\epsilon^2)}$.

Occupation of respiratory equipment — Most people that died from Covid-19 required respiratory equipment until their death for a period of length τ_r and we assumed this period τ_r

to be constant. If γ people out of all that require respiratory equipment die, we can estimate the daily occupation of respiratory equipment $r(t)$ by summing over the past τ_r days of newly seriously sick persons per day (SSPs), which are related to the daily deaths shifted by the typical time T from being diagnosed as seriously sick until death. For that, we divide the sum of deaths over the past τ_r days by γ to extrapolate to active SSPs at time t and hence required respiratory equipment

$$\begin{aligned} r(t) &= \frac{1}{\gamma}\int_{t-\tau_r}^{t} \mathrm{SSP}(t')\,\mathrm{d}t' = \frac{1}{\gamma}\int_{t+T-\tau_r}^{t+T} d(t')\,\mathrm{d}t' \\ &= \frac{D(t+T)-D(t+T-\tau_r)}{\gamma}. \end{aligned} \tag{A8}$$

The number of required respiratory machines $r(t)$ attains its maximum at time $t_{\mathrm{r,max}} = t_{\mathrm{d,max}} - T + \tau_r/2$ and thus the peak number of required respiratory machines is $r_{\mathrm{max}} = r(t_{\mathrm{r,max}}) = (D_{\mathrm{total}}/\gamma)\,\mathrm{erf}(\tau_r/2w_d)$, where D_{total} is the total number of deceased people. This peak r_{max} increases with larger occupation times of respiratory machines τ_r, larger total number of fatalities D_{total} and narrower GM widths w_d. Flatten the curve!

Percentiles of infection numbers — From $t_{d,\mathrm{max}}$ and w_d we can estimate dates at which the number of daily infected people had reduced to the level of $\eta \in [0,1]$ of its maximum value. These times denoted as T_η are given by

$$T_\eta = t_{d,\mathrm{max}} - \tau - w_d\sqrt{\ln\eta} \tag{A9}$$

For $\eta = 1\%$ and $\eta = 1‰$ these times are explicitly given, employing the typical delay time $\tau \approx 10$ days (A6), by

$$\begin{aligned} T_{1\%} &= t_{d,\mathrm{max}} - 10 + 2.146\,w_d, \\ T_{1‰} &= t_{d,\mathrm{max}} - 10 + 2.628\,w_d \end{aligned} \tag{A10}$$

The corresponding dates are listed in Fig. 4. It is also possible to estimate dates for which less than a certain η of the total population remains infected and potentially dangerous to initiate another outbreak. This time is given by

$$T_\eta^* = t_{i,\mathrm{max}} + w_i\,\mathrm{erf}^{-1}\left(1 - \frac{\eta f}{2D_{\mathrm{total}}/\mathrm{MP}}\right), \tag{A11}$$

where $D_{\mathrm{total}}/\mathrm{MP}$ had been tabulated (Fig. 4), erf^{-1} is the inverse error function, and f defined by (A6) may be approximated by the value mentioned there. In this expression $t_{i,\mathrm{max}}$ and w_i can also be approximated by using the values calculated from fatalities, as described already. For $\eta = 10^{-6}$ (one per million inhabitants), and a typical $w_i = 10$, $T_y^* - t_{i,\mathrm{max}} \approx 46$ days for $D_{\mathrm{total}}/\mathrm{MP} = 100$.

Doubling times — Doubling times, here denoted by k, are used to characterize the strength of an exponential growth process, independent of the exponential amplitude. A doubling time quantifies the time span required for the exponential to double (or, up to convention, to have doubled). Assuming a purely exponential growth, both $d(t) = d_{\mathrm{max}}\exp(\nu t)$ and $D(t) = d(t)/\nu$ increase mono-exponentially with time, and the doubling time k is a constant, $k = \nu^{-1}\ln 2$, while $\nu = d'(t)/d(t) = D'(t)/D(t)$. For the GM the doubling time based on $d(t)$ is thus given by

$$k(t) = \frac{k(0)}{1 - t/t_{\mathrm{d,max}}}, \quad k(0) = \frac{(\ln\sqrt{2})w_d^2}{t_{\mathrm{d,max}}}, \tag{A12}$$

while the doubling time based on $D(t)$ is given by $k(t) = (\ln 2)/[k\ln D(t)/dt] = (\ln 2)D(t)/d(t)$. It is thus easy to calculate two versions of doubling times with the GM parameters at hand, using either daily or total measures, which differ if the growth is not ideally exponential. While doubling times are convenient as they alter only weakly during exponential growth, they are difficult to extract from data directly without applying smoothing procedures that differ from publication to publication, and they are not uniquely defined. For this reason we do not recommend to proceed with an analysis on reported doubling times, as done in [8], unless the raw data is unavailable.

[1] Organisation for economic co-operation and development (oecd). URL https://data.oecd.org.

[2] R. S. Abhari, M. Marini, and N. Chokani. COVID-19 epidemic in Switzerland: Growth prediction and containment strategy using artificial intelligence and big data. doi: https://doi.org/10.1101/2020.03.30.20047472.

[3] D. Adam. Special report: The simulations driving the world's response to COVID-19. *Nature*, 2020. doi: doi:10.1038/d41586-020-01003-6. available online.

[4] M. an der Heiden and U. Buchholz. Modellierung von beispielszenarien an der sars-cov-2 epidemie 2020 in deutschland. 2020. doi: 10.25646/6571.2. in German.

[5] 72 different github contributors as of April 2. Json time-series of coronavirus cases (confirmed, deaths and recovered) per country - updated daily. URL https://pomber.github.io/covid19.

[6] X. Fu, Q. Ying, T. Zeng, and Y. Wang. Simulating and forecasting the cumulative confirmed cases of SARS-CoV-2 in China by boltzmann function-based regression analyses. *J. Infection*, 2020. doi: https://doi.org/10.1016/j.jinf.2020.02.019. available online.

[7] R. Ospinac F. A. G. Almeidad G. C. Duarte-Filhod I. C. L. Souzae G. L. Vasconcelosa, A. M. S. Macedob. Modelling fatality curves of covid-19 and the effectiveness of intervention strategies. URL https://doi.org/10.1101/2020.04.02.20051557.

[8] R. Schlickeiser and F. Schlickeiser. A gaussian model for the time development of the Sars-Cov-2 corona pandemic disease. predictions for germany made on march 30, 2020. 2020. doi: https://doi.org/10.1101/2020.03.31.20048942.

[9] H. Stevens. Why outbreaks like coronavirus spread exponentially, and how to flatten the curve. *Washington Post*. URL https://www.washingtonpost.com/graphics/2020/world/corona-simulator/.

[10] X. Yang, Y. Yu, J. Xu, H. Shu, J. Xia, H. Liu, Y. Wu, L. Zhang, Z. Yu, M. Fang, T. Yu, Y. Wang, S. Pan, X. Zou, S. Yuan, and Y. Shang. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *The Lancet*, 2020. doi: https://doi.org/10.1016/S2213-2600(20)30079-5. available online.

α_3 code	Country	w_i	$t_{i,\text{max}}$	i_{max}	I_{total}	$I_{\text{total}}/\text{MP}$
BRA	Brazil	11.1 ± 0.2	04-03 $\pm$ 8	800 ± 350	16000 ± 7000	78 ± 35
CHL	Chile	11.7 ± 0.2	04-02 $\pm$ 7	321 ± 33	6600 ± 800	370 ± 45
GBR	Great Britain	20.0 ± 0.7	04-27 $\pm$ 14	—	—	—
JPN	Japan	38.0 ± 2.3	04-27 $\pm$ 22	195 ± 90	13000 ± 7000	104 ± 54
SAU	Saudi Arabia	15.4 ± 0.6	04-02 $\pm$ 13	140 ± 23	3900 ± 700	120 ± 24
SRB	Serbia	10.3 ± 0.3	04-01 $\pm$ 8	113 ± 50	2100 ± 1000	290 ± 130
PAK	Pakistan	10.3 ± 0.3	03-28 $\pm$ 11	170 ± 80	3200 ± 1600	16 ± 8
PER	Peru	16.8 ± 1.3	04-09 $\pm$ 28	148 ± 64	4400 ± 2300	140 ± 70
POL	Poland	15.0 ± 0.4	04-07 $\pm$ 10	340 ± 50	9100 ± 1500	240 ± 40
ROU	Romania	18.9 ± 1.3	04-19 $\pm$ 26	690 ± 90	23200 ± 4600	1200 ± 250
USA	United States	14.8 ± 0.2	04-14 $\pm$ 5	—	—	—

TABLE I: GM parameters w_i, $t_{i,\text{max}}$, and i_{max} for those countries, for which sufficient data about infections, but insufficient data about fatalities is available to us, as of April 2. The coefficient i_{max} we could not extract from the existing data with an error less than 100%.