DOUBLE POWER LAW FOR COVID-19: PREDICTION OF NEW CASES AND DEATH RATES IN ITALY AND SPAIN

Vladimir A. Osherovich[1]

Joseph Fainberg[2]

Lev Z. Osherovich[3]

[1] NASA/GSFC/CUA Code 673, Greenbelt, MD 20771, US

Email: Vladimir.Osherovich@gmail.com

[2] NASA/GSFC Emeritus, Code 673, Greenbelt, MD 20771, US

Email: fainberg@jhu.edu

[3] Versant Ventures (Switzerland) GmbH, Aeschenvorstadt 36, Basel 4051 Switzerland

Email: losherovich@versantventures.com

1. Abstract

The novel corona virus SARS-CoV-2 appeared at the end of 2019, spreading rapidly and causing a severe respiratory syndrome (COVID-19) with high mortality (2-5%). Until a vaccine or therapy is found, the most effective method of prophylaxis has been to minimize transmission via rigorous social distancing and seclusion of all but essential workers. Such measures, implemented at different times and to varying degrees world-wide, have reduced the rate of transmission compared with early phases of the pandemic, resulting in "flattening of the curve" followed by a gradual reduction in mortality after >6 weeks of rigorous social distancing measures. The cost of rigorous social distancing has been seen in radically reduced economic activity, job losses, disruption of schooling and social institutions. A key question facing policy makers and individuals is when to resume normal economic and social activity in the face of persistent community transmission of SARS-CoV-2. To help address this question, we have developed a model that accurately describes the entire transmission and mortality curves in Italy and Spain, two hard-hit countries that have maintained severe social distancing measures for over 2 months. Our model quantitatively describes the rapid rise and slow decay of new cases and deaths observed under stringent social distancing (the "long tail" effect). We predict that even when social distancing is rigorously maintained, the number of COVID-19 deaths after peak mortality may be 2 – 3 times larger than the total number of deaths up to the peak. Our model has important policy implications for countries currently debating how to ease social distancing measures.

2. Methods

To model the infection rate and death rate curves for Italy and Spain, we propose a Double Power Law (DPL) equation with 4 parameters. In this model, we treat the spread of SARS-CoV-2 as a multiscale process with two time scales: one time constant (m1) defines the growth rate and another time constant (m3) defines the relaxation process that results from effective social distancing measures. m1 is on the order of a few days while m3 is at least one order of magnitude larger (30-60 days). The other two constants (m2 and m4) are non-dimensional and relate to the powers of growth and relaxation phases. The formula we propose for our model is given in Equation 1:

$$Y = \frac{(t/m1)^{m2}}{1+(t/m3)^{m4}} \qquad (1)$$

where the evolution function Y can be for new cases per day (NC) or death rate (DR). Best fit values for the parameters m1, m2, m3 and m4 were determined by least-squares fit to the available data and are shown in Table 1. Primary data is drawn from https://www.worldometers.info/coronavirus/

3. Results

We have analyzed data from Italy and Spain, two countries that experienced rapid initial growth of NC and instituted rigorous social distancing early during the pandemic. Importantly, the primary data span times before and after NC and DR peaks, thus providing a complete picture of COVID-19 dynamics before and during social distancing measures.

Figures 1a and 1b show the NC and DR for Italy. The black circles are daily values averaged over 3 days for the period from February 22, 2020 to April 19, 2020. Figure 1c shows the total number of deaths (without 3-day averaging) summed up to each day as black open circles. In Figures 1a and 1b, the red line is the resulting theoretical fit of the 3-day average of the Italian daily data. In Figures1a and 1b, the red line represents the fitting for the interval February 22. 2020 to April 19 2020 extended beyond the interval. The regression coefficient (R) for this fit in Figure 1a is R=0.98 and for Figure 1b is R=0.99 with values of the parameters shown in Table 1. For both NC and DR, we have found that data up to April 19, 2020 are sufficient to find the parameters m1, m2, m3 and m4 (Table 1). In Table 1, t_{tmax} refers to the beginning of fitting time (in Figures 1 and 2). Note that these fitting start times (as calendar days) are different for NC and DR fits. For NC, peak days and peak values are presented in Table 2. Prospective data gathered after the construction of the model are shown as blue squares (labeled as "Observations") and follow closely the theoretically extrapolated curve.

In Figure 1c, the total (cumulative) number of deaths (TD) is shown by the red line which has been extended out to 100 days from the beginning. The black circles are the sum of reported daily deaths up to that date. Root mean square deviation of observed data is given as sigma on the plots and in Table 1.

The solid red curve in Figures 1a,b,c is the theoretical extrapolation of the fit predicted by the DPL equation (1). The projection of DR out to 100 days after February 22, 2020 shows an asymmetric fit or long tail. For Italy, the total number of deaths at 100 days is 3.68 times larger than the total deaths at the peak of DR according to Table 3. The open circles in Figure 1c are the sums of the daily death rates (without 3-day averaging) up to that date. As in Figures 1a and 1b, we mark prospective data values obtained after April 19, 2020 that were not utilized in the curve fitting. We find that the regression coefficient between TD observations (open circles) and our model fit (red line) is 0.99.

The results for Spain (Figure 2) are similar to those of Italy for both DR and for total death TD. Figure 2b shows a somewhat faster growth of DR at the beginning (March 1, 2020); m2 obtained for Italy is smaller

compared to m2 for Spain. The relaxation for Spain appears shorter since its value for m4-m2 is slightly larger than the value obtained for Italy. The projection of the ratio of total deaths at day 100 to the total deaths at the day of the DR peak for Spain is 3.69 which is slightly larger than for Italy. Tables 1, 2 and 3 contain all relevant parameters for Spain. Comparing m3 from Table 1 with t_{max} for NC and DR, we have found that m3 is comparable to t_{max} for Italy and Spain.

We have also noticed weekly oscillations of NC data from Italy, Spain and Germany, which we illustrate in Figure 3. The sudden surge of prospective Spanish NC data (seen as a bump on the tail of Figure 2a) after April 19 is discussed in Appendix 1 together with the weekly oscillation phenomenon.

4. Discussion

In this paper we have suggested a Double Power Law in the form of Equation (1) as the way to describe the whole curve for the time evolution of the new cases and death rate in one country. For the beginning of the process (small t), Equation (1) has an asymptotic power law ~ $(t/m1)^{m2}$. For large times, $Y \sim t^{m2-m4} m3^{m4} m1^{-m2}$, thus Y decreases if m4 > m2. In this case, for m2>0 for small t, Y increases as a power law and for large t, Y decreases also with a power law. This is the reason we refer to Equation (1) as a Double Power Law. Often the intial growth phase of COVID-19 case load is described as exponential, but our fits show that the growth more accurately follows a power law, as does the decrease of Y after the peak time. Based on observed and prospective data, m2 is positive and controls the growth of the fit. The value of m2-m4 is negative and thus controls the relaxation phase of the curve after the peak, corresponding to effective implementation of social distancing measures.

We show that our equation fits new cases and death rates for Italy and Spain with regression coefficients of 0.98 to 0.99. The high degree of adherence of prospective data (blue squares) to theoretical projection of DPL curves for NC and DR validates the DPL model and allows extrapolation into the future,

providing that social distancing measures remain in place. Such projections can serve as a baseline to the impact of relaxation of social distancing conditions on new cases and deaths. Additionally, we have demonstrated the phenomena of weekly oscillations in NC for Italy, Spain and Germany (Appendix 1). The nature of these oscillations will be addressed in future work (manuscript in preparation).

Other researchers have attempted to model dynamics of the SARS-CoV-2 pandemic, focusing on theoretical and practical factors governing transmission rates. Social distancing as the means to reduce the transmission of the corona virus has been modelled in a recent influential paper by Ferguson et al. 2020,. Kissler *et al.* 2020 model the effects of seasonal transmission, acquired immunity and non-pharmaceutical interventions such as social distancing on viral reproductive number R_0. Blasius (2020) applies a single power law to the intial growth phase of SARS-CoV-2, citing early mathematical models of smallpox transmission by Daniel Bernoulli (1766). Other than our Equation (1), we have not found other studies that apply a Double Power Law for modeling COVID-19 or any other epidemiological processes.

Critiques of the use of power law models in biological phenomena focus on lack of statistical validation for such approaches and lack of underlying mechanistic knowledge about the parameters that drive these models (Stumpf and Porter, 2012). While we acknowledge complex and as-yet unclear biological underpinnings for SARS-CoV-2 transmission, there is strong statistical validation of the DPL equation for COVID-19 incidence and mortality. Firstly, we have reproduced the observed death rates and total deaths with high accuracy ($R>0.98$) for two countries up to April 19, 2020 using Equation (1). Secondly, our model prospectively anticipates data gathered after April 19, 2020 which were not included in intial model fitting. More work is needed to derive a full understanding of the biological factors that drive the growth and relaxation parameters in Equation (1).

Interestingly, our DPL model displays an asymmetric curve similar to Planck's black body radiation law. At low frequencies, Planck's distribution follows the Raleigh-Jeans power law, while at large frequencies, the relaxation or decrease of Planck's curve is exponential. Like Planck's law, our DPL model for

COVID-19 NC and DR displays distinct power law behavior for growth and decay, with transition from one power law to another closely consistent with observations near the peak.

The implications of our model for public policy are that premature slackening of social distancing measures will allow resumption of growth phase, resulting in increased deaths above the fit projections, which are based on continued social distancing. Critically, if social distancing measures are withdrawn before NC falls below sigma, our model predicts that growth in NC will resume in a m1-dominated fashion, although possibly with different parameters. Due to the observed long (>5 day) incubation period of the virus and the typically ~10-14 day period of progression from exposure and death, the resumption of rapid (m1-dominated) growth will not be apparent until at least two weeks after complete relaxation of social distancing measures. As many countries are now relaxing their social distancing measures, we expect that the DPL model could be used to parametrize and compare the effect of different policies and practices on NC and DR. Likewise, when new vaccines and therapies are developed and implemented, changes in the parameters of the DPL model can be used measure the impact of interventions on NC and DR. We intend to use the methodology developed in our paper to study its relevance to COVID-19 data from the United States, Great Britain and Germany and to develop biological hypotheses for parameters of the DPL equation (work in preparation).

5. Appendix 1.

The maximum of NC and DR according to Formula (1) is described as

$$Y_{max} = \frac{(t_{max}/m1)^{m2}}{1+(t_{max}/m3)^{m4}} \qquad (2)$$

where t_{max} is the time when Y_{max} occurs. At the time of the maximum, the first derivative of Y must be zero. The last condition allows us to express t_{max} in terms of the constants m1, m2, m3 and m4

$$t_{max} = m3 * (m4/(m2-1))^{\left(\frac{-1}{m4}\right)} \qquad (3)$$

Formula (3) allows us to express m2 as a function of m3, m4 and t_{max}. Thus, theoretically instead of m1, m2, m3, and m4 as fitting parameters, the set of 4 numbers Y_{max}, t_{max} , m1 and m3 can be used. It is important to recognize that once Y_{max} and t_{max} are known, together with the whole growth part of the curve up to the peak, then the relaxation part of the curve is completely determined.

6. Appendix 2.

For new cases presented in Figure 1a for Italy, observational data points (black circles) form a double peak. In fact, there are 7-day oscillations around the smooth double power fit particularly evident during the relaxation phase of the NC curves. For DR (Figure 1b), these oscillations are also present. In Figure 3, we show NC normalized by the DPL fit for three cases. The black curve is for Italy, the red is for Germany and the blue is for Spain. To produce the German data, we verified that NC for Germany also fits well by our DPL Equation (1). We also observed weekly oscillations in data for the NC and DR for other countries including the United States as well as pooled global data. Examining this in detail is beyond the scope of this paper.

7. Acknowledgements.

The authors are grateful to Elena Romm at NIH/National Institute of Allergy and Infectious Diseases and Jordan Fainberg for discussions and suggestions.

8. References.

Blasius, B. https://www.researchgate.net/publication/340394479_Power-law_distribution_in_the_number_of_confirmed_COVID-19_cases, (2020)

Bernoulli, D., "Essai d'une nouvelle analyse de la mortalité causée par la petite vérole" . Math. Phys. Acad. Roy. Sci., Paris, (1766)

Ferguson, N.M. et al., "Impact of nonpharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand", DOI:https://doi.org/10.255b1/77482 (2020).

Kissler S.M. et al. *"*Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period," DOI: 10.1126/science.abb5793 (2020)

Stumpf, M.P.H and M.A. Porter. "Critical Truths About Power Laws" DOI: 10.1126/science.1216142 (2012)

9. Figures

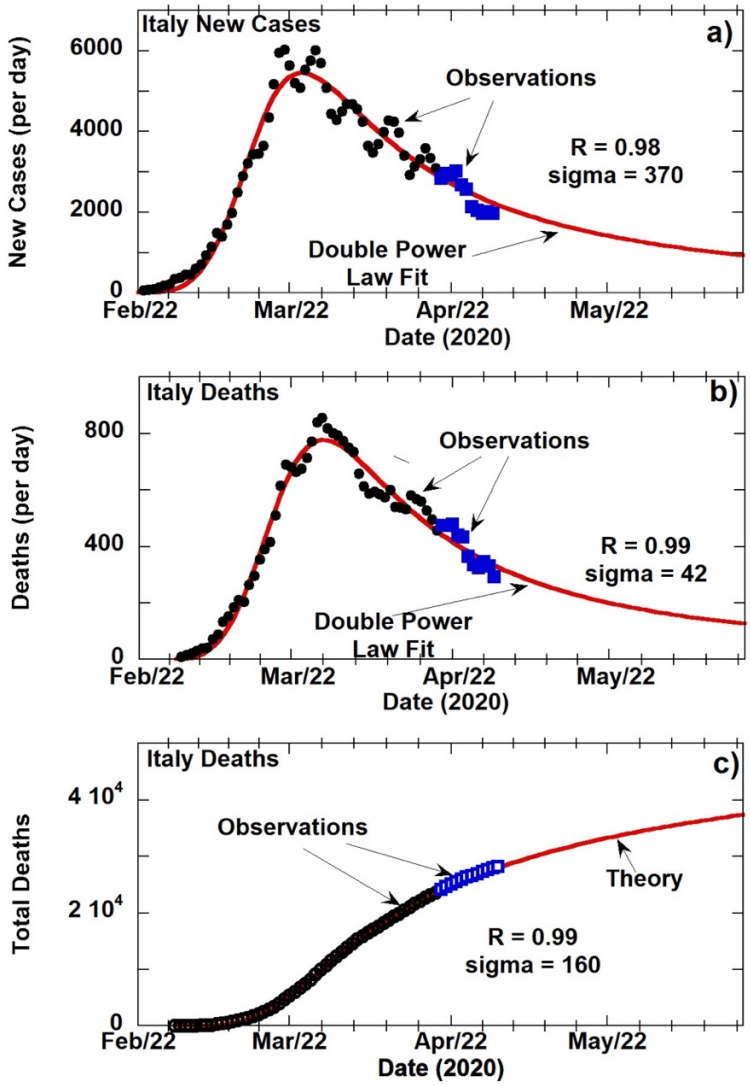

Figure 1. **a)** COVID-19 new cases per day in Italy with a. 3-day averaging starting from February 22, 2020; **b)** Death rate (deaths per day) in Italy starting at February 28, 2020; **c)** Cumulative death numbers from February 28, 2020 in Italy. Black circles in a) and b) are observed values up till April 19, 2020. These data have been used for the Double Power Law fits (red lines). The blue squares are values observed after April 19, 2020 and are not used in the fits. In c) open circles represent total death numbers summed up without 3-day averaging for that date with blue squares for data after April 19, 2020.

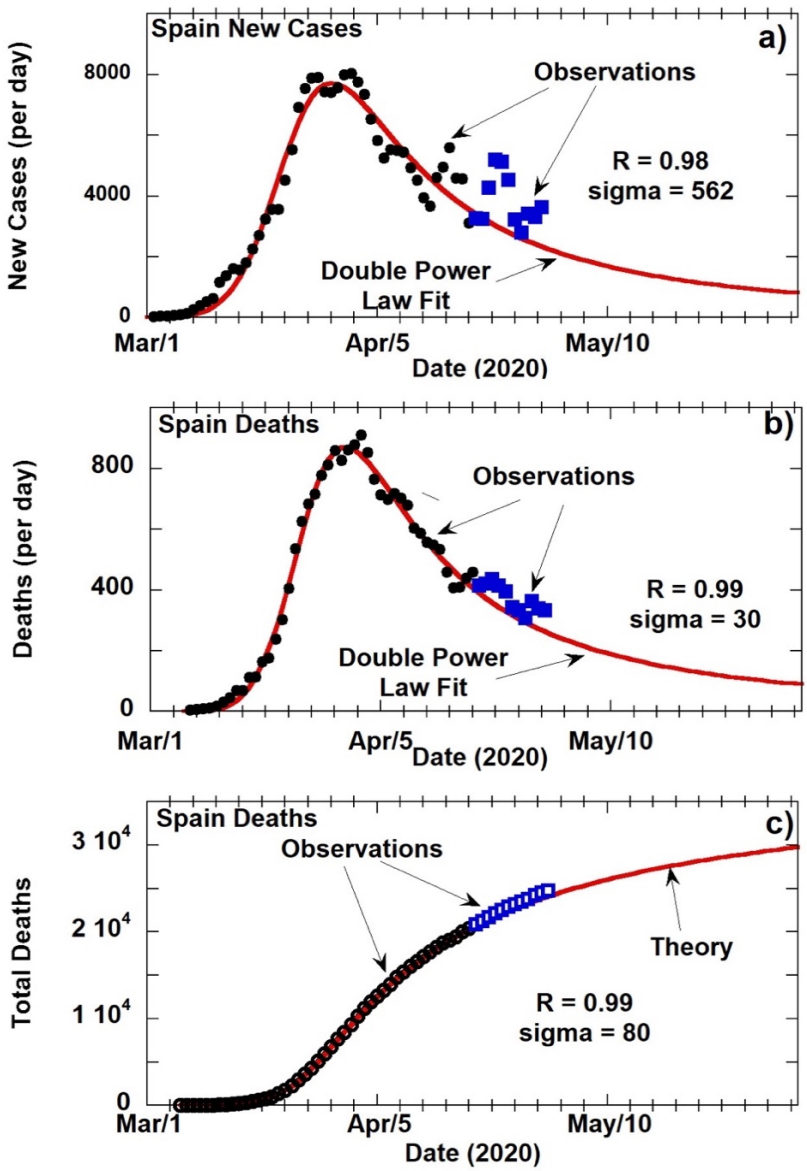

Figure 2. a) COVID-19 new cases per day in Spain with 3-day averaging starting from March 1, 2020. The blue squares represent data after April 19, 2020 and represent the large 7-day oscillation for Spain seen in Figure 3; **b)** Death rate (deaths per day) in Spain; **c)** Total death number from March 6, 2020 in Spain. Descriptions same as in Figure 1.

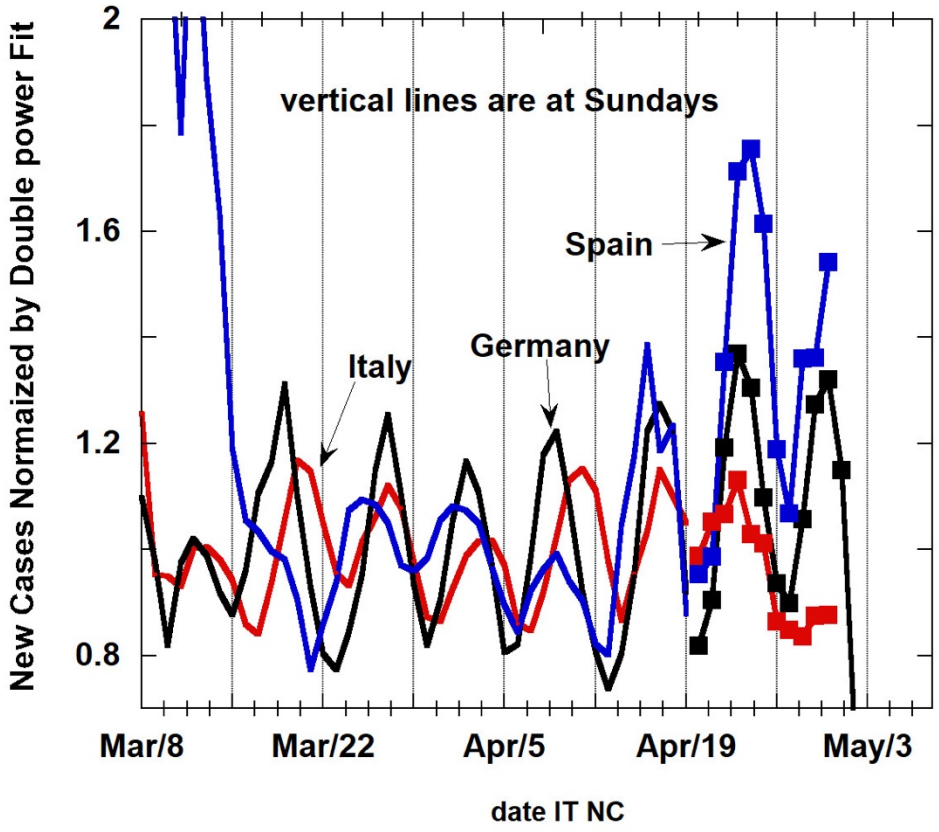

Figure 3. Weekly oscillations of new cases in Italy (black), in Spain (blue) and in Germany (red) with Sundays shown by vertical lines. The 3-day averages are normalized by the Double Power Law fit values. The squares are observed values after April 19, 2020 which was the end dates for the fits. Spain shows large weekly oscillations in late April. The minima of these oscillations appear near weekends.

10. Tables

Table 1. Fit Parameters for New Cases and Death Rates

Parameter	Italy	Italy	Spain	Spain
	NC	DR	NC	DR
m1	2.49 day	2.21 day	2.79 day	3.76 day
m2	3.83	2.97	4.29	4.11
m3	27.59 day	26.0 day	26.1 day	22.86 day
m4	5.45	4.58	6.35	6.15
R	0.98	0.99	0.98	0.99
sigma	370 per day	42 per day	562 per day	30 per day
t_{max}	32.3 day	29.25 day	29.1 day	25.61 day
Y_{max}	5463/day	778/day	7723/day	877/day

Table 2. Fit Characteristics for New Cases (2020)

Data Fit	Italy Feb 22 to Apr 19	Spain Mar 1 to Apr 19
Peak day	March 24	March 29
Peak value	5455 per day	7722 per day

Table 3. Fit Characteristics for Death Rates and Total Deaths (2020)

Data Fit	Italy Feb 28 to Apr 19	Spain Mar 6 to Apr 19
Peak day	March 28	March 31
Peak value	777 deaths/day	871 deaths/day
Sum: Day 0 to Peak day	9722 deaths up to Peak	8417 deaths up to Peak
Sum: Peak Day to day 100	27082 deaths after Peak	22646 deaths after Peak