Covid-19 Case Fitting Curves

Esmerindo Bernardes[1,*]

[1]*São Carlos Institute of Physics, University of São Paulo, São Carlos, SP, Brasil*

(Dated: April 27, 2020)

Based on data from Covid-19 cases from China, a universal fitting curve is presented. Only five parameters are required to be determined by a given Covid-19 case data over a number of days. Under the special condition of reliable mass testing, this curve can indicate the flattening of a given case data after its inflection point is reached.

The hyperbolic tangent function,

$$Z_N(n) = A\tanh(an - b) + C, \tag{1}$$

has two desirable features of any Covid-19 case data over the number of days n: it is flat at both beginning and end, and has one inflection point. Magnitude and steepness of the accumulated cases over N days are given by parameters A (amplitude) and a (slope), respectively.

Besides China, which has a complete case data, other countries having total cases above or near China are analyzed. US, Spain, Italy, France, Germany and Iran seem to be close of their inflection points. UK and Turkey seem to be not so close of their inflection points. Nonetheless, it is important to say that despite the (weak) evidences presented below, any prediction here is not to be taken for granted. Also, this is a simple exercise of reconversion, since I am not an expert in the Statistics Science. See Refs. [1, 4] for more accurate fitting schemes.

Figure 1 shows the case curves for the complete Covid-19 case data from China. The whole data seems to have a subseries described by the curve Z_{21} (the violet curve), obtained using the first 21 days, which indicates a (local) flattening of the cases over two weeks (around 53,700 cases). In spite of this curve Z_{21} being a false lead in searching for the case flattening, since we know the whole history, it has all features of a typical (complete) case curve. In fact, a close analysis in this subseries shows an inflection point around $N = 14$. Also, while the nearest case curves $Z_{N\leq 21}$ approach Z_{21} in an oscillatory way, all case curves $Z_{N\geq 23}$ approach monotonically the global case curve Z_{70} (the black curve in Figure 1), obtained using the first 70 days in the case data from China, which indicates a flattening around 81,000 cases. Therefore, the (black) curve Z_{70} shown in Figure 1 is the lower limit for all curves $Z_{N\geq 22}$ and has its inflection point around $N = 21$. All curves $Z_{N\geq 29}$ converge monotonically and rapidly to Z_{70}, giving a 10 days prediction for the (global) flattening. Best parameters for these two curves are shown in Table I.

We have learned that the case curves Z_N obtained after the inflection point converge monotonically, in an increasing speed, to the global case curve. Of course, mass testing is presumed. Figure 2 shows the root mean square (rms) deviations of each case curve $Z_{N\geq 22}$, confirming the monotonic convergence.

Some questions. What did happen at day 22? Was it an increase in the total tests? Or is this jump in the number of cases from day 21 to the day 22 (see the solid green circles in Figure 1) a typical feature of the global case curve around its inflection point (where the rate of change is maximum)?

Let us try to analyze case data from other countries based on what we have learned from the Chinese case data, under the model function (1). Figure 3 shows the case curves over the last five days for the USA case data. Since these curves are ordered, with the last day (blue) curve at the top, and quite distant from each other at the flattening region, it shows (probably) that the inflection point was not reached yet, but it could be close. The case curves from Spain, Italy, Germany, UK and Turkey have these same features (see Figures 4–8). Russia and Brazil seem to be far away from their inflection points (see Figures 12–13).

Jumps. The case data from France shown in Figure 9 has two jumps: at day 40 to 41 and at day 53 to 54 (see the solid green circles). Two subseries is something new. The case data from UK seems to have one jump at day 45 to 46 (see the solid green circles Figure 7), but it can be diminished by the growth in case data.

Flattening. France and Iran have case curves indicating a possible flattening (the last curve is a lower limit), as shown in Figures 9 and 10. Reinforcing that, for Iran only, Figure 11 shows that the root mean square (rms) deviation seems to have reached its maximum. However, these flattenings could be false leads due to oscillations in the case data.

Residuals[1]. There are growing oscillations in the residuals from US, Spain, Italy and Germany case data. Iran has the most regular oscillations in their residuals, as shown in Figure 14. These oscillations superpose a modulation in the case data, which can not be accounted by the hyperbolic tangent in the model function (1). Perhaps a q-deformed hyperbolic tangent could delivery a better response, using the deformation parameter to control its steepness.

What about the planet Earth? Figure 15 shows the world case data. It seems to be far away from the inflection point. Better stay at home.

* sousa@ifsc.usp.br

[1] Residual: the difference between reported and predicted cases, using the last day fitted curve.

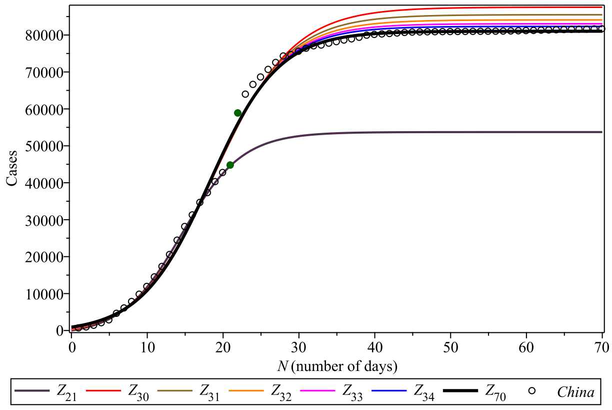

FIG. 1. Covid-19 case curves for the data from China. Day 1 is January, 22, with 571 cases. Parameters are shown in Table I.

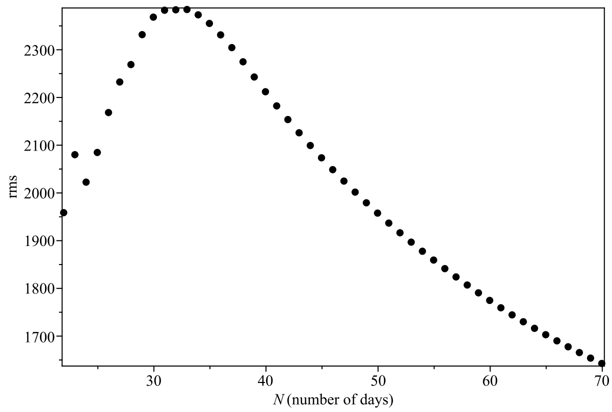

FIG. 2. Root mean square (rms) deviation for the Covid-19 case curves $Z_{N\geq 22}$ for the data from China.

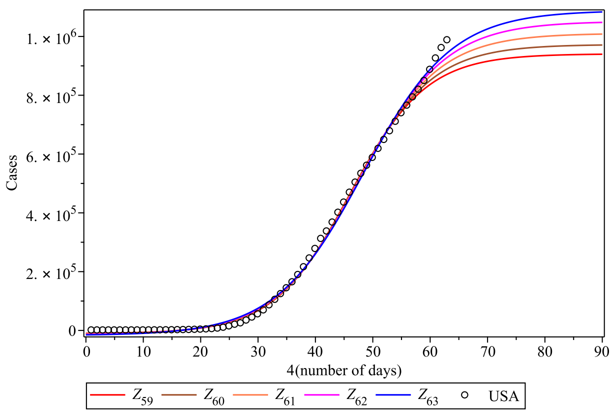

FIG. 3. Covid-19 case curves for the data from USA. Day 1 is February, 15.

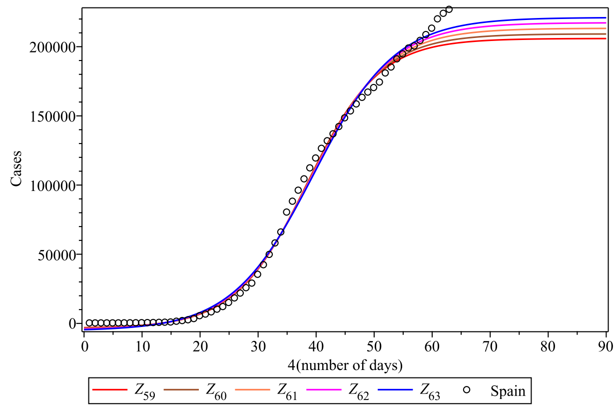

FIG. 4. Covid-19 case curves for the data from Spain. Day 1 is February, 23.

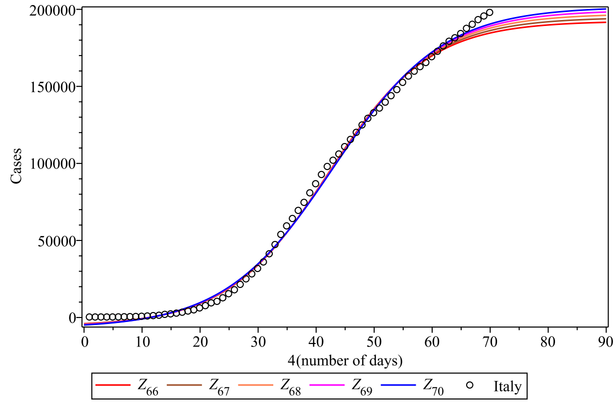

FIG. 5. Covid-19 case curves for the data from Italy. Day 1 is February, 20.

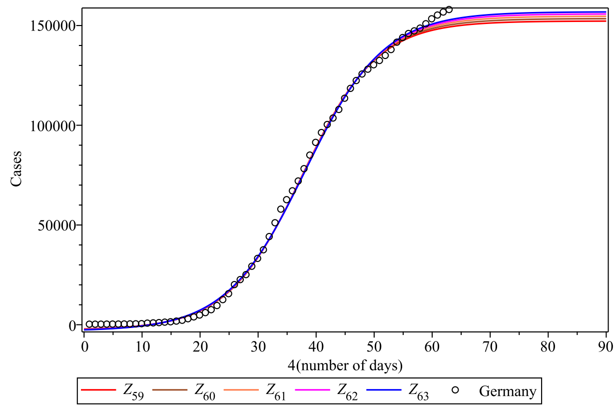

FIG. 6. Covid-19 case curves for the data from Germany. Day 1 is February, 24.

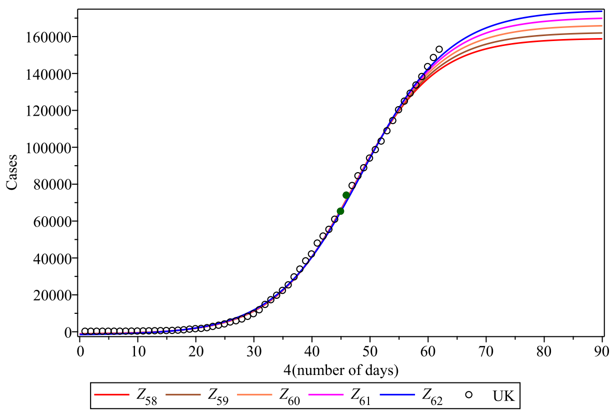

FIG. 7. Covid-19 case curves for the data from UK. Day 1 is February, 22.

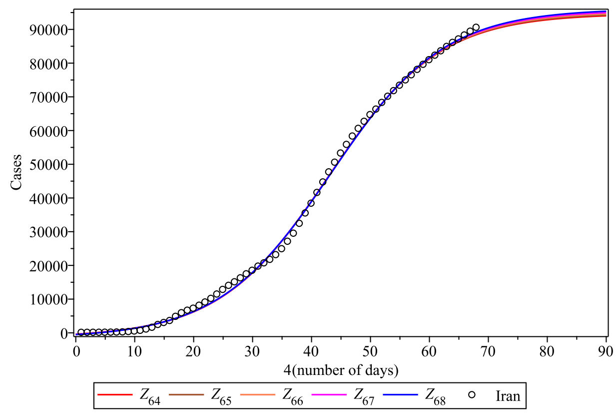

FIG. 10. Covid-19 case curves for the data from Iran. Day 1 is February, 19.

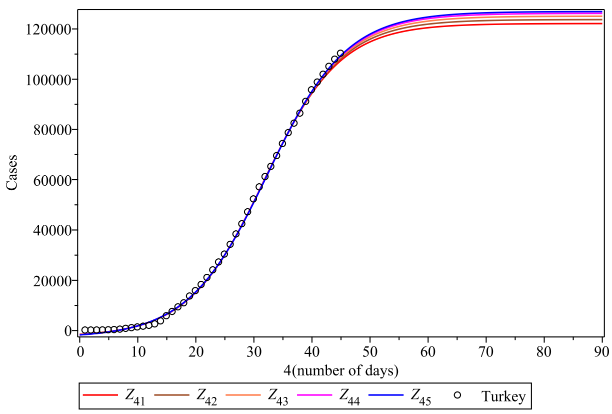

FIG. 8. Covid-19 case curves for the data from TU. Day 1 is February, 22.

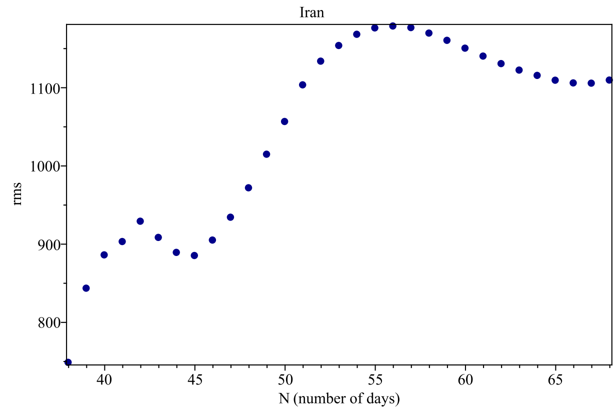

FIG. 11. Root mean square (rms) deviation for the case data from Iran.

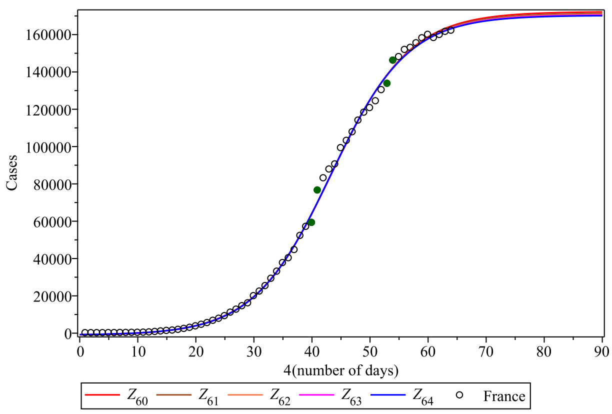

FIG. 9. Covid-19 case curves for the data from France. Day 1 is February, 23.

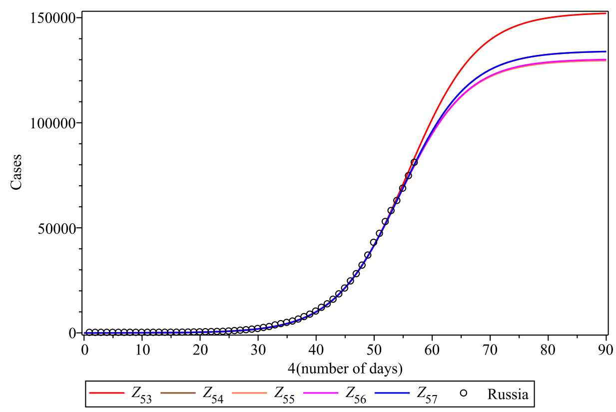

FIG. 12. Covid-19 case curves for the data from Russia.

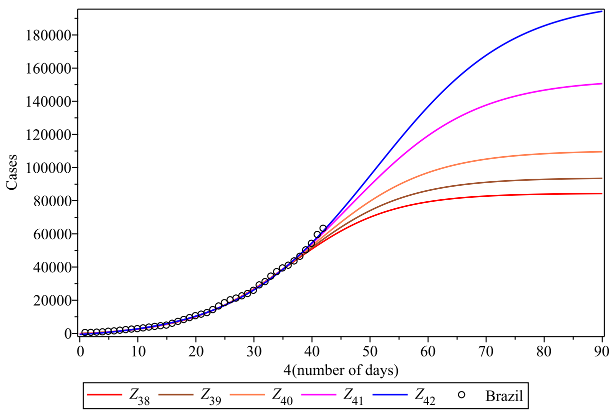

FIG. 13. Covid-19 case curves for the data from Brazil.

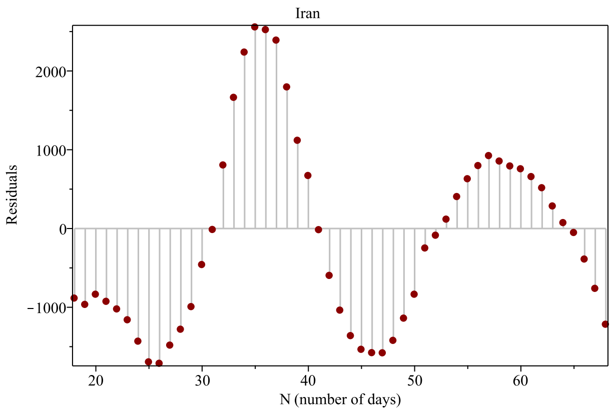

FIG. 14. Iran residuals: the difference between reported and predicted cases, using the last day fitted curve.

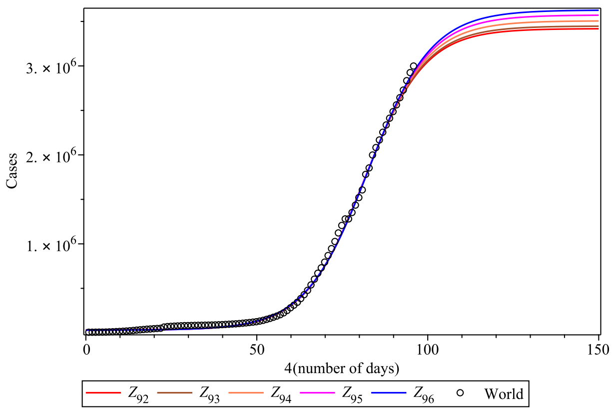

FIG. 15. Covid-19 case curves for the data from planet Earth.

Parameters $\{A, a, b, C\}$ appearing in the model function (1) were obtained inside Maple and Mathematica by a (local) non-linear fitting minimizing the least-squares error (both worksheet and notebook can be found here). All case data points are weighted equally. Case data from every country are available at Worldometer, Coronavirus.

	A	a	b	C
Z_{21}	27537.6	0.12651	3.683	26181.1
Z_{70}	40770.9	0.11020	3.625	40252.4

TABLE I. Parameters for the Covid-19 case curves for the data from China.

[1] S. Maltezos, Parametrization Model Motivated from Physical Processes for Studying the Spread of COVID-19 Epidemic. arXiv:2004.05992v2 [physics.soc-ph] (2020).
[2] P. Girardi *et al.*, Robust inference for nonlinear regression models from the Tsallis score: application to Covid-19 contagion in Italy. arXiv:2004.03187v2 [stat.AP] (2020).
[3] M. A. V. Arias, Using generalized logistics regression to forecast population infected by Covid-19. arXiv:2004.02406v1 [q-bio.PE] (2020).
[4] J. Kumar and K. P. S. S. Hembram, Epidemiological study of novel coronavirus (COVID-19). arXiv:2003.11376v1 [q-bio.PE] (2020).