Italian Covid-19 epidemic @ 14 March 2020: logistic and others

Vincenzo Fiorentini
Dipartimento di Fisica, Università di Cagliari, Cittadella Universitaria, Monserrato, I-09042 Cagliari, Italy

Update **14 March 2020**; time evolution looks logistic (or in any event slowly saturating) on all variables, with decreasing uncertainties on saturation predictions. The focus now shifts onto other fit models and their properties. Here I look at the Gompertz model and its current parameters.

I report time evolution estimates of the Covid-19 italian epidemic, hypothesizing a logistic behavior. The data are nation-wide up to March 14 (Protezione Civile, `https://bit.ly/2UbpPzt`). The previous reports are at `https://bit.ly/2W6vs4u` as well as in posts at `https://bit.ly/2QaFQEy`.

I then analyze the data using a Gompertz function, and provide its parameters as determined today. These are also compared to similar estimates of the Hubei deaths data.

The early-stage onset of saturated behavior is quite settled for all quantities. The exponentials can now be rejected based on statistical indicators, and are not considered in this report (except in Fig.1). The predicted points of inflection and, especially, saturation values are still extremely uncertain. Fig.1 visualizes the data. Fit parameters as well as real and simulated data are available at `https://bit.ly/2W6vs4u`. A simplified python code is available. For work on individual Italian regions, I point out again Edoardo Gorini's work at `https://bit.ly/2IHbc1p`. S. Miscetti, N. Tomassetti, and O. Maragò at `https://bit.ly/2QaFQEy` also provided estimates similar to mine.

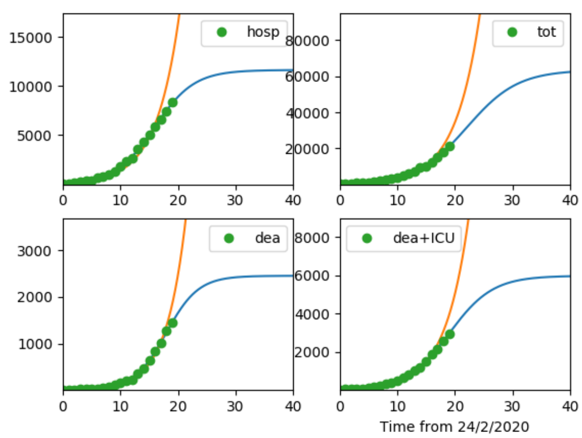

FIG. 1. Exponential and logistic behavior (lin-lin scale) for (top-down, left-right) hospitalized, total infected, deceased, ICU+deceased.

Fig.1 and the statistics in Table I point fairly clearly to a good logistic fit. Based on current data, the saturations for deaths+ICU, deaths, hospitalized, total infected are 6000, 2500, 11630, 63000, compared to 6200, 3000, 10500, 48000 yesterday, so still quite fluctuating. The inflection points for the same quantities are 19, 18, 16, 22 days from 24 February, i.e. around roughly March 15, nearly stable compared to the last values. Again cautionary flag (which may be for the better in a sense) is the partial inconsistency of the different estimates. Extrapolated final lethality (deaths/positives at the end of the process) is currently still large at 4%, and the death rate for hospitalized patients still huge, at 22%.

TABLE I. Statistical comparison (13 March 2020) for logistic (L) and Gompertz (G).

	S-p	paired-S-p	χ^2	χ^2-p	R^2
Deaths					
L	0.988	0.432	9.546	0.963	0.999
G	0.974	0.287	54.882	0.0	0.998
ICU+deaths					
L	0.994	0.581	9.083	0.972	1.0
G	0.994	0.642	11.222	0.916	1.0

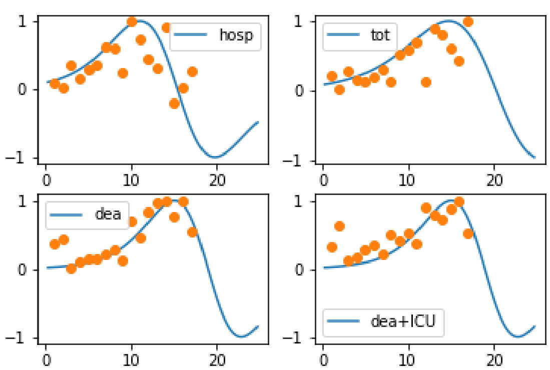

FIG. 2. Second derivatives calculated numerically for the data (dots) using a 3-point formula and smoothing by averaging over ±2 points, and for the fit (line) with the 3-points formula.

At the inflection points, the second derivatives of the various quantities should become negative. In Fig.2 I report them for the logistic fit and the data. There is a lot of noise and fluctuation, but the rough behavior seems to be in place. It is obviously impossible to find the inflection using the data directly.

As mentioned yesterday, we are treading into the domain of quantitative predictions or analysis, and it is important to experiment with other possible fitting models. At least at this stage in the game, these models do not necessarily correspond to dynamical models (although some do) and are purely fitting assumptions. Another important point is that we are now dealing with a one-population problem whereby the data are dominated, for better or worse, by Lombardia. It remains to be seen if another exponential outbreak, due to other regions becoming critical, may superimpose on the forthcoming plateau.

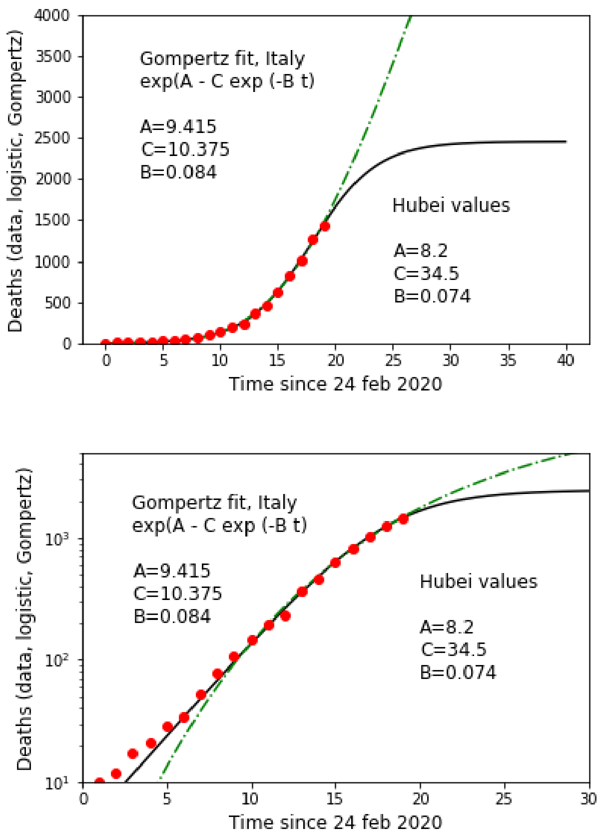

FIG. 3. Gompertz and logistic fit to the deaths data up to March 14, 2020 (linear, top; log-lin, bottom). The fit parameters are reported, along with those of the Hubei data (not shown) by F. Ricci-Tersenghi, `https://bit.ly/2TSYbs2`.

In any event, I have studied the Gompertz model function $f(t, A, B, C) = \exp(A - C \exp(-Bt))$, an exponential with decreasing (in fact, exponentially decreasing) rate (see `https://bit.ly/38PCVaH` , `https://bit.ly/2TOsIHo`). F. Ricci-Tersenghi showed recently on `https://bit.ly/2TSYbs2` that Gompertz fits satisfactorily the Hubei deaths data, whereas the logistic provides a decent but imperfect fit. A merit of this function is that its derivative ("the peak") is asymmetric as is expected of the real distribution.

The result of the fit is shown in Fig.3. Clearly the Gompertz model has generally a long way to go before plateauing off; if this were the correct model with the current parameters, we would be in for a long wait for the epidemic to end. However, for the moment, aside from the fit statistics still in favor of the logistic (Table I), the extreme sensitivity of the Gompertz to changes in the rate makes it even harder to make stable predictions than with the logistic. The current deaths saturation would be 12500 with inflection at 28 days, i.e. on March 23, which is potentially plausible given our relatively slow response compared to China. The total infected saturating at over 3 million are less plausible, especially coupled with 69 days for the inflection point, obviously at odds with the deaths peak. Also inconsistent with deaths, and further suggesting an inaccurate prediction in this regime, is the deaths+ICU number, saturating at 60000 and with inflection at 60 days (whereas one would expect deaths and deaths+ICU to saturate to the same value).

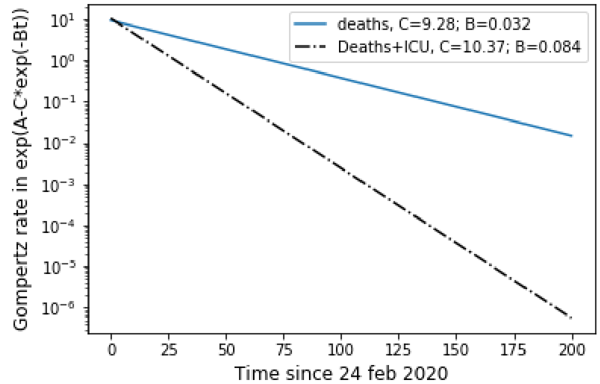

FIG. 4. Gompertz rate for deaths and logistic fit to the deaths data up to March 14, 2020.

One reason for the sensitivity is clearly the exponentially changing rate, which is plotted in Fig.4 for deaths and deaths+ICU in log-lin scale. The decay parameter is similar to the Hubei case, but the leading coefficient is smaller, causing the larger rise and extended duration. Whether this is realistic or not remains to be seen.