Italian Covid-19 epidemic @ 18 March 2020: logistic and Gompertz

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

Update **18 March 2020**. I compare logistic, Gompertz and their derivatives to monitor the possible appearance of "the peak".

I study the time evolution estimates of the Covid-19 italian epidemic using nation-wide data up to March 18 (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`.

The onset of non-exponential behavior has been settled for a few days. An extrapolation at long times is now strongly dependent on the choice of the fitting function. Here I compare the logistic and Gompertz functions. It is appropriate to state unambiguously that this is purely a fitting exercise, and does not account for epidemiological considerations. In particular the main reason to use the Gompertz function is just that it seems to fit well the data for the Hubei outbreak (and the logistic is a reasonable step function). It is expected that extrapolations will improve as data amass and become more reliable, and that, for example, the rates of change of the cumulant quantities can be better understood.

Both functions track closely the data and are statistically acceptable (unlike yesterday the Gompertz is slightly favored; for brevity I am omitting the indicators); we cannot yet discard either hypothesis. In order to increase the sensitivity to short-time changes, I have been reporting first derivatives for a couple of days now, and today I show them for a few proxies of the epidemic in Figs.1 to 4 (in lieu of the proxies themselves and their fits).

Today the rates appear to be branching off onto the Gompertz peak rather than bending down on the logistic peak. Taken at face value, this would imply a death rate due to increase over the next days, and peak at 700/day at the end of March. It is still early to make even semi-quantitative estimates with any degree of certainty, but for the moment the analysis does not bode well for the near future.

The cautiously optimistic outlook based on the data of the last couple of days has been dashed by a death rate increase of 34% nation-wide over that of the three previous days, and worse, a 50% increase in Lombardia, accounting by itself for the whole difference compared to recent days. Now this number is quite insane; as somebody quipped a few days ago, "Come on, it's an epidemic, not a plane crash." Clearly these data bear witness to the collapse of local medical structures, as well as problems of data reporting. For another example, yesterday we had 140 declared healed, whereas today we had 1050. Both numbers taken individually are clearly unrealistic and cast doubts on the way they are aggregated in the reports.

Table I reports the current saturations and inflections for the last four days. Saturations stay basically in the same ballpark as yesterday for death count (for total infected, the Gompertz saturation is still insane at 480k, but down from 650k, 1.3 million, 3.5 million in the previous days). Inflection times are stable for both functions (very unstable for Gompertz totals infected).

TABLE I. Extrapolations with logistic and Gompertz for deaths and total infected (data 18 March 2020). The four values per entry are for the last four days including today.

	Deaths	Total
	Saturation	
L	3283/4118/4543/5500	73250/67811/64710/67500
G	19870/26375/23409/31000	3.8 M/1.4 M/0.65 M/0.48 M
	Inflection date	
L	15/3, 16/3, 16/3, 16/3	18/3, 18/3, 17/3, 17/3
G	27/3, 30/3, 29/3, 30/3	05/5, 21/4, 15/4

As a side note, I have taken some flak of various nature for my posting of March 17 (apparently, reading abilities are not a common gift). I nevertheless stand by the idea that monitoring the rate of increase is a valuable exercise. Also, the use of Gompertz function has been criticized. I think its use is justified as it amounts essentially to an exponential with a variable rate. Data fitting with so called segmented exponentials is fairly common (`https://bit.ly/3a4GACV` in medical statistics: exponentials are assigned different constant rates in different time slices, with their values changing in time in some chosen fashion. Gompertz simply assumes that the rate is exponential, and for the present case, decreasing. This appears to be a decent approximation for the Hubei data. For its very form, Gompertz is very sensitive at large times (for example, using a short polynomial expansion of the rate the estimates are quite different – and more optimistic).

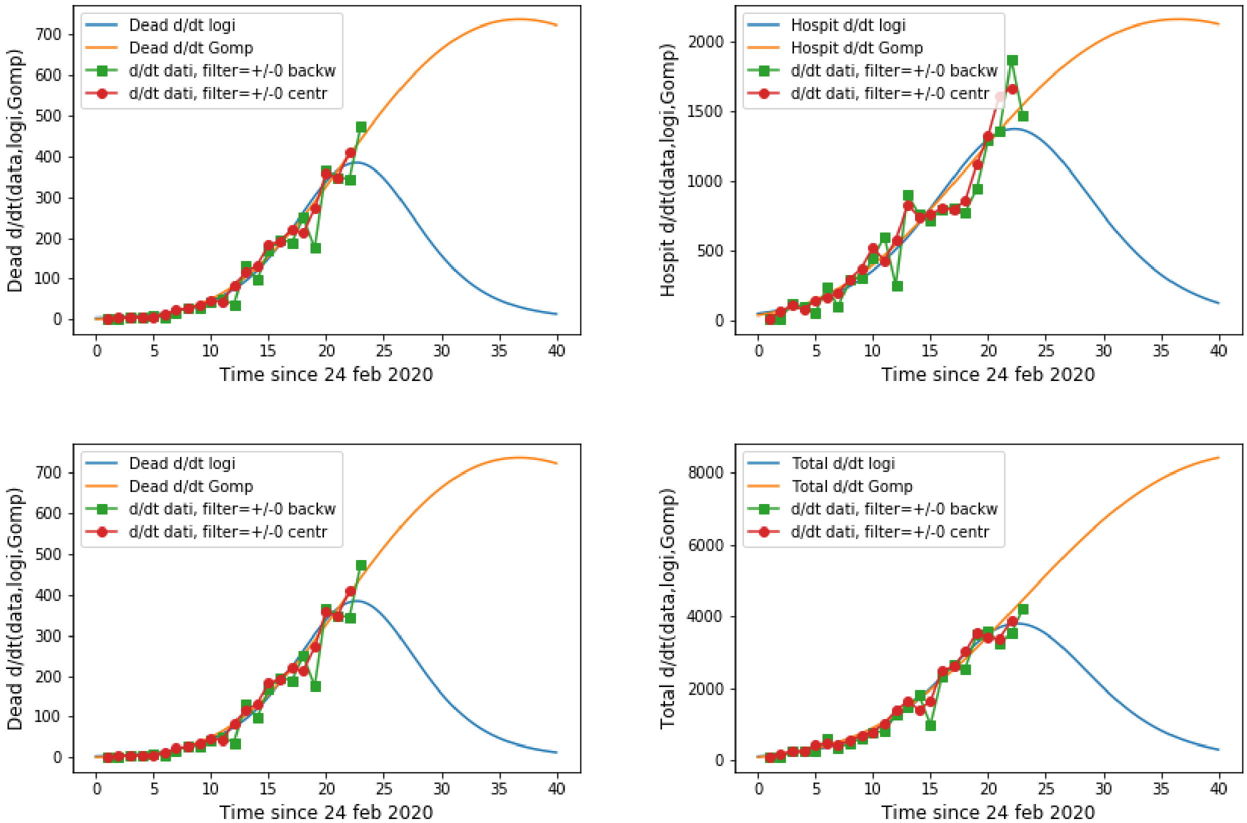

FIG. 1. Derivatives of logistic and Gompertz fits, and backward (squares) and centered (circles) derivative of the data for various proxies (as indicated in the legends).