Modelling of COVID-19 mortality

Stanislav Volkov*

May 4, 2020

1 The model

Let $t = 1, 2, \ldots, T$ be the day since the first covid-related death in a country, T being the final date for which data are available. Denote by N_t the *total* number of people who died by date t. The data for N_t are widely available, for example from `https://www.worldometers.info/coronavirus`. Let $k_t = N_t/N_{t-1} \geq 1$ be the relative daily increase in the total deaths. The purpose of this note is to find a good temporal model for the growth of N_t.

Since N_t is non-decreasing, it is clear that $k_t \geq 1$. Therefore, it is convenient to assume that

$$k_t = 1 + \exp(h(t))$$

for some function $h(t)$.

As it turns out, when assuming that $h(t)$ is a polynomial, a very good fit is obtained if

$$h(t) = a + bt^2$$

for some parameters a, b, which are to be estimated.

Assuming $N_0 = 1$, we have

$$N_t = \prod_{i=1}^{t} k_i = \prod_{i=1}^{t} \left(1 + e^{h(i)}\right) \approx \exp\left\{\sum_{i=1}^{t} e^{h(i)}\right\} \approx \exp\left\{\int_0^t e^{a+bu^2}\,\mathrm{d}u\right\} =: f(t)$$

where we assumed that k_t is quite close to 1. Since $\frac{\mathrm{d}\tilde{N}(s)}{\mathrm{d}s}$ gives the daily number of deaths, we might compute

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \exp\left\{a + bt^2 + \int_0^t e^{a+bu^2}\,\mathrm{d}u\right\}$$

*Centre for Mathematical Sciences, Lund University, Box 118 SE-22100, Lund, Sweden

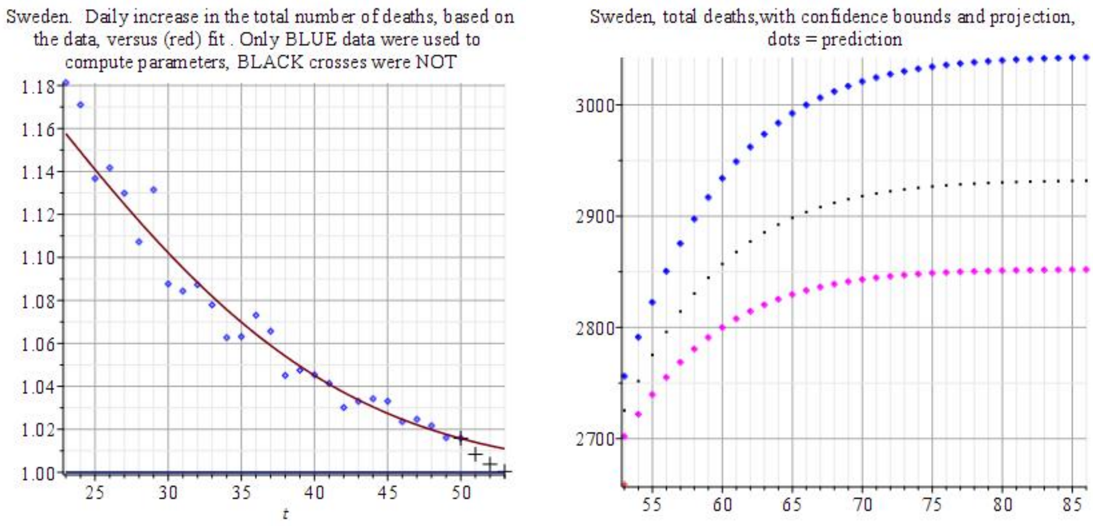

Figure 1: Sweden

In particular, the peak number of deaths is achieved when

$$0 = \frac{\mathrm{d}^2 f}{\mathrm{d}t^2} = \frac{\mathrm{d}f}{\mathrm{d}t} \cdot \exp\left\{2bt + e^{a+bt^2}\right\}$$

i.e., when

$$t = t_{\max} = -\frac{e^{-\frac{1}{2}\mathrm{LambertW}\left(-\frac{e^{2a}}{2b}\right)+a}}{2b}. \quad (1)$$

Finally, since there are various policy changes which arrive with time (lockdown, etc.) that can affect the parameters of the model, it's reasonable to use more recent data in our case we typically use the consecutive three-four weeks.

2 Sweden

The data for Sweden are taken from the official Swedish health ministry site `https://www.folkhalsomyndigheten.se/smitts887kydd-beredskap/utbrott/aktuella-utbrott/covid-19/bekraftade-fall-i-sverige/`. We also ignore the last 4 days, as the data for last few days always get substantially updated later due to delays in reporting.

We obtain that k_t takes values

$$1.181, 1.171, 1.137, 1.142, 1.130, 1.107, 1.131, 1.088, 1.084, 1.087, 1.078, 1.063, 1.063, 1.073,$$
$$1.066, 1.045, 1.048, 1.045, 1.041, 1.030, 1.033, 1.034, 1.033, 1.024, 1.025, 1.022, 1.016, 1.016,$$

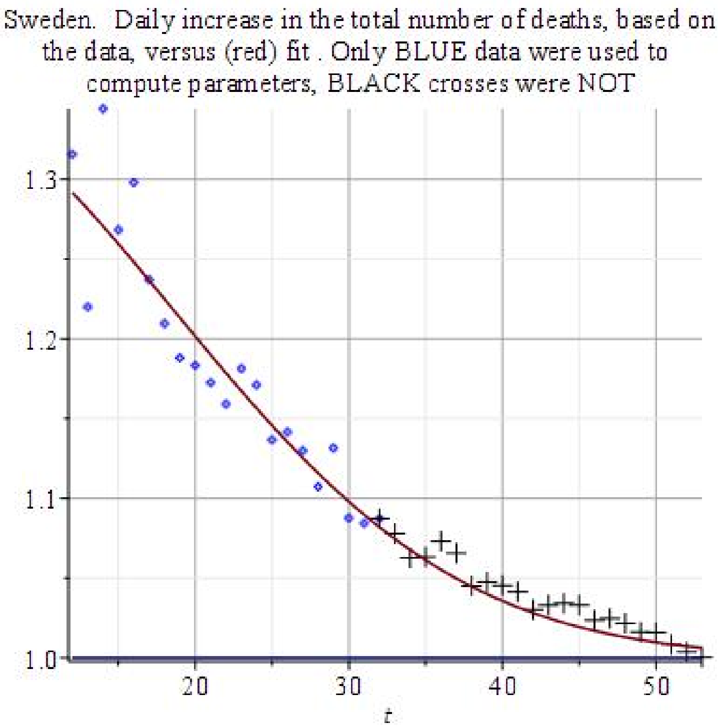

Figure 2: Predictive power of the model

Using OLS for 28 data points in the equation

$$\ln(k_t - 1) = a + bt^2,$$

we get

$$a = -1.23\ (0.06), \quad b = -0.00117\ (0.00004), \qquad R^2 = 0.974$$

where standard deviations of the estimates are given in the brackets. Using (1), we guess that the peak number of deaths in Sweden was around 2 weeks ago (mid April).

We present the graph which shows goodness of fit and future prediction on Figure 3. To judge the predictive power of the model, we estimated the parameters using only the data for week -4,-5,-6, got the prediction for weeks -3,-2,-1, and compared them with the actual data which we have for these weeks on Figure 2.

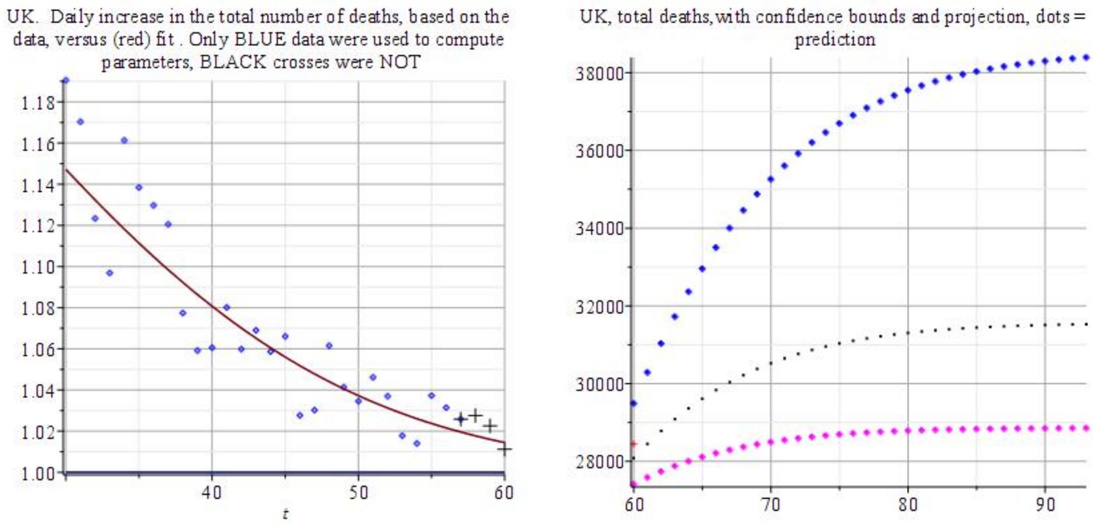

Figure 3: United Kingdom

3 UK

In case of the UK, we use the data from the Worldometer site. We get the values of k_t as

$$1.191, 1.170, 1.123, 1.097, 1.161, 1.138, 1.130, 1.121, 1.077, 1.059, 1.061, 1.080, 1.060, 1.069,$$
$$1.059, 1.066, 1.028, 1.030, 1.062, 1.041, 1.035, 1.046, 1.037, 1.018, 1.014, 1.037, 1.031, 1.026,$$

and thus

$$a = -1.14\ (0.18), \quad b = -0.00086\ (0.00008), \qquad R^2 = 0.798$$

using the latest 4 weeks of data. The peak number of deaths also seem to have occurred around 2-3 weeks ago.