A model to forecast the evolution of the number of COVID-19 symptomatic patiens after drastic isolation measures

Luis Alvarez

University of Las Palmas of Gran Canaria, Spain

email: lalvarez@ulpgc.es

March 24, 2020

1 Introduction

In this work we propose a model to forecast the evolution of the number of COVID-19 symptomatic patients after drastic isolation measures. Our main assumptions in this case are :

1. The evolution of the number of symptomatic patients, $N(t)$, grows at exponential rate (that we name a), during a period of time t_0, then, after drastic isolation measures it grows at a residual exponential rate (that we name b) until a certain time $t_1 > t_0$, and after t_1 we assume that no new patients appear.

2. Each new contaminated patients is added to $N(t)$ according to the probability distribution of the incubation period.

3. The data of contaminated patients provided by the countries can be considered that it concern mostly to symptomatic patients. This a reasonable assumption in the countries where tests are done only to patients which show some kind of symptoms.

Regarding the distribution of the incubation period, Lauer et al. in [1], using the data of 181 patients approximate the distribution of the incubation period as a log-normal distribution. The cumulative distribution function of this log-normal is given by

$$F(t) = \begin{cases} \int_0^t \frac{e^{-\frac{(\log x - \mu)^2}{2\sigma^2}}}{x\sigma\sqrt{2\pi}} dx & if \quad t > 0 \\ 0 & \text{otherwise} \end{cases}$$

with $\mu = 1.621$ and $\sigma = 0.418$.

Accordingly with these assumptions the evolution of patients $N(t)$ is given by

$$N(t) = \sum_{k=0}^{t_0} Ce^{ak}(e^a - 1)F(t-k) + \sum_{k=t_0+1}^{t_1} Ce^{at_0}e^{b(k-k_0)}(e^b - 1)F(t-k) \tag{1}$$

where $C > 0$ is a constant. For simplicity the model is formulated in a discrete way using a time step of 1.

2 Parameters estimation

We point out that the parameters of the above model are given by C, a, b, t_0 and t_1.

Given a dataset, $D(t)$, of the number of symptomatic patients across time for a region, we fix the parameters by minimizing the quadratic mean error

$$Error(C, a, b, t_0, t_1, \tilde{t}) = \frac{1}{(t_{\max} - t_{\min} + 1)} \sum_{t=t_{\min}}^{t=t_{\max}} \left(D(t) - N(t + \tilde{t})\right)^2$$

where $\tilde{t}$ is the translation of $N(t)$ to fit $D(t)$. $[t_{\min}, t_{\max}]$ is the range of values we use to fit the parameters of the model. We use a basic gradient descent type method to adjust the parameters.

Table 1: Parameters adjustment for the country datasets. The parameter $\tilde{t}$ represents the displacement in days with respect to the first day taking for the analysis (January 20, 2020).

	China	China	Italy	Spain
$t_{\min}$	January 27	January 27	February 28	March 8
$t_{\max}$	March 21	February 16	March 21	March 22
C	2847	3377	620	4686
a	0.157762	0.159805	0.162020	0.147345
b	0.031299	0.044838	0.019462	0.004581
t_0	18	16	27	16
t_1	31	30	32	25
$\tilde{t}$	4	3	-29	-45
$Error$	$1.992165e \times 10^6$	6.951823×10^5	9.901376×10^5	8.511217×10^4

3 Experiments

We will use the data set of the evolution of contaminated patients in China, Italy and Spain taking from the web page `https://ourworldindata.org/coronavirus-source-data`. Concerning China, in February 13, it seems that China changed the way they counted the number of contaminated people and then an artificial jump appears in the evolution. To correct this error, we multiply the previous values by a factor $M = 1.255$ in order to smooth the transition. In table 1, we show the results obtained for the different countries. For comparison purposes, in the case of China we use 2 intervals $[t_{\min}, t_{\max}]$ to adjust the parameters. We observe that in all cases, the values of the exponential grow a in the case of no-isolation measure is quite similar and, as expected, the residual exponential grow b is much smaller. In Fig. 1, 2, 3 we compare the expected evolution accordingly with the model and the actual data. In the case of China, using a large interval of times to estimate the parameters the expected evolution is not far from the actual one all along the times. Using a shorter period of times at the beginning of the full period (to check the ability of the method to forecast future evolutions), we observe that the final number of patients is larger than the actual one, but still reasonable. In the case of Italy and Spain the results are quite reasonable and the model provides a forecast of the evolution. In spite of we can expect (as in China) a significant error in the final estimation, we think that it could be good enough to help to understand how the expected evolution could works.
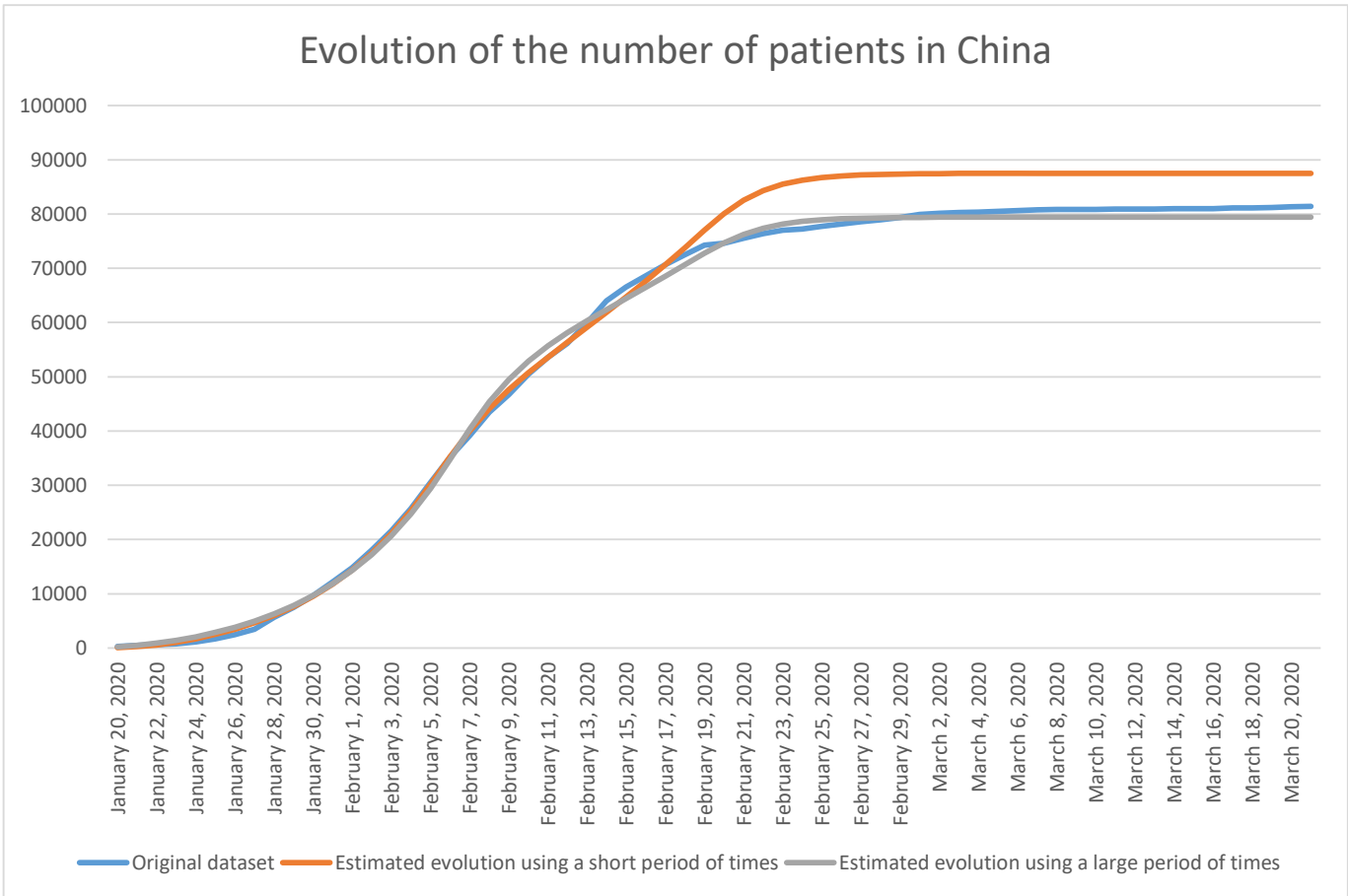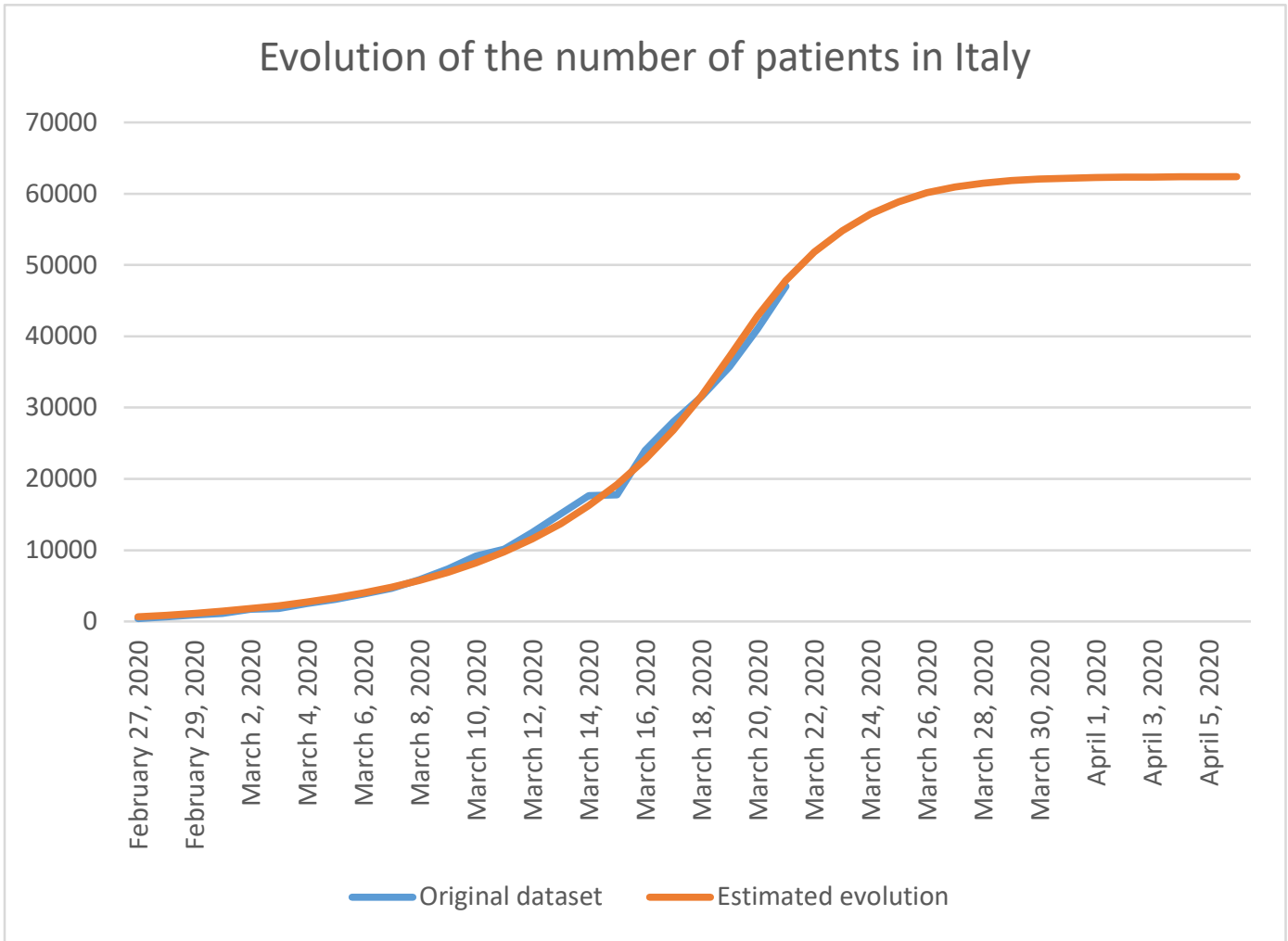
Figure 1: Evolution of patients in China

Figure 2: Evolution of patients in Italy, using the model, we can forecast the evolution in the future
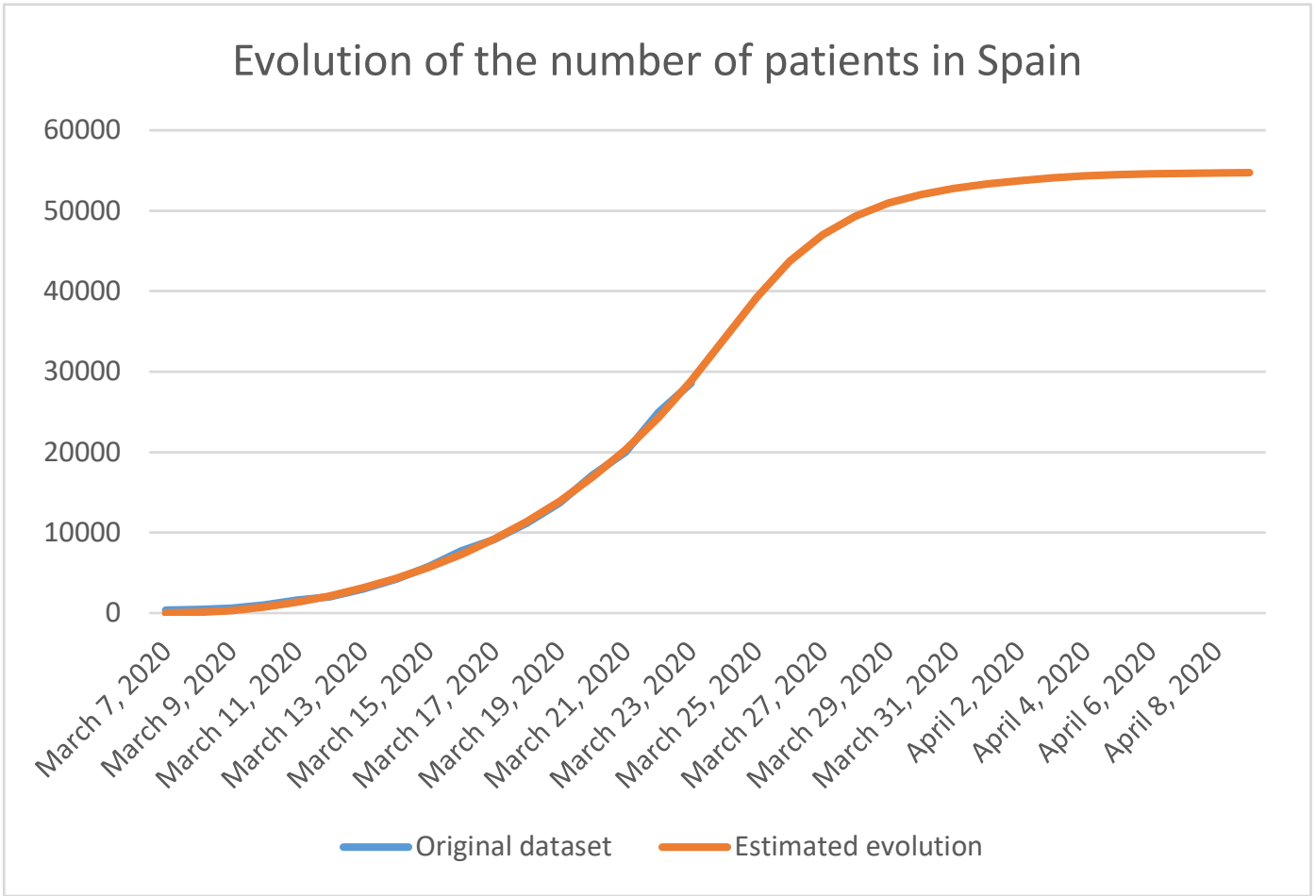
Figure 3: Evolution of patients in Spain, using the model, we can forecast the evolution in the future

References

[1] Stephen A. Lauer and MS, PhD * and Kyra H. Grantz, BA * and Qifang Bi, MHS and Forrest K. Jones and MPH and Qulu Zheng, MHS and Hannah R. Meredith, PhD and Andrew S. Azman and PhD and Nicholas G. Reich PhD and Justin Lessler, PhD *"The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application"*. Annals of Internal Medicine, 2020. DOI: 10.7326/M20-0504