

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/339943580>

# Analysis and forecast of COVID-19 spreading. !!! UPDATE !!!

Preprint · March 2020

CITATIONS

0

READS

1,432

2 authors:



[Duccio Fanelli](#)

University of Florence

248 PUBLICATIONS 1,959 CITATIONS

[SEE PROFILE](#)



[Francesco Piazza](#)

Université d'Orléans

97 PUBLICATIONS 1,080 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Workshop on Self-organized patterns on complex networks and the brain [View project](#)



Systems with long-range interactions [View project](#)

# Analysis and forecast of COVID-19 spreading

Duccio Fanelli<sup>1</sup>, Francesco Piazza<sup>2,3</sup>

<sup>1</sup>*Dipartimento di Fisica e Astronomia, Università di Firenze,  
INFN and CSDC, Via Sansone 1, 50019 Sesto Fiorentino, Firenze, Italy\**

<sup>2</sup>*Centre de Biophysique Moléculaire (CBM),  
CNRS-UPR 4301, Rue C. Sadron, 45071 Orléans, France and*

<sup>3</sup> *Université d'Orléans, Château de la Source, 45071 Orléans Cedex, France†*

This document is the update to the project “Analysis and forecast of COVID-19 spreading” uploaded on March 15<sup>th</sup> 2020. The project is live on ResearchGate at the address <https://www.researchgate.net/project/Analysis-and-forecast-of-COVID-19-spreading>.

---

\* Duccio.Fanelli@unifi.it

† Francesco.Piazza@cnrs-orleans.fr

## CONTENTS

I. Theory: basic SIRD model and variant with non-constant infection rate	2
II. Analysis of the outbreak in South Korea	3
III. Updated fitting for the spreading of the epidemics in Italy.	5
Acknowledgments	5
References	5

### I. THEORY: BASIC SIRD MODEL AND VARIANT WITH NON-CONSTANT INFECTION RATE

Within the simplest model of the evolution of an epidemic outbreak, people can be divided into different classes (species). In the susceptible (S), infected (I), recovered (R), dead (D) scheme (SIRD), any individual in the fraction of the overall population that will eventually get sick belongs to one of the aforementioned classes. Let  $S_0$  be the size of the initial population of susceptible people. The mean-field kinetics of the SIRD epidemic evolution is described by the following system of differential equations

$$\begin{aligned}
 \frac{dS}{dt} &= -rSI \\
 \frac{dI}{dt} &= rSI - (a + d)I \\
 \frac{dR}{dt} &= aI \\
 \frac{dD}{dt} &= dI
 \end{aligned}
 \tag{1}$$

with initial condition  $[S(t_0), I(t_0), R(t_0), D(t_0)] = [S_0, I_0, R_0, D_0]$  for some initial time  $t_0$ . The parameter  $r$  is the infection rate, i.e. the probability per unit time that a susceptible individual contract the disease when entering in contact with an infected person. The parameters  $a$  and  $d$  denote, respectively, the recovery and death rates.

Many countries have introduced harsh, draconian containment measures, with the aim of reducing the infectivity rate. To model this effect, one can consider a modified version of the SIRD model, where the infection rate  $r$  is let vary with time. As a simple example, given that the

containment measures became law at time  $t^*$ , one can take

$$r(t) = \begin{cases} r_0 & \text{for } t \leq t^* \\ r_0(1 - \alpha)e^{-(t-t^*)/\Delta t} + \alpha r_0 & \text{for } t > t^* \end{cases} \quad (2)$$

where  $r_0$  is the rate estimated from the fit to the data in some early-time interval that is supposedly unaffected by the lockdown, and  $\alpha \in [0, 1]$  gauges the asymptotic reduction of the infection rate afforded by the containment measures. The parameter  $\Delta t$  represents the average time to wait for the effects of the lockdown to be visible.

## II. ANALYSIS OF THE OUTBREAK IN SOUTH KOREA

South Korea was among the first countries to implement massive testing (of the order of 10,000 tests/day during the peak phase) and strong containment measures. A simultaneous fit of the CSSE data in the period 22/01/2020 - 14/03/2020 is shown in Fig. 1 (top panels), while a simultaneous fit to the data with the modified SIRD model with non-constant infection rate given by Eq. (2) is illustrated in Fig. 1 (bottom panels). The SIRD best-fit parameters are reported for both cases in Table I. For the modified SIRD model, we find from the fit

$$t_0 = 40.8 \pm 0.1 \quad \text{days}$$

$$\Delta t = 3.32 \pm 0.05 \quad \text{days}$$

$$\alpha = (4 \pm 2) \times 10^{-3}$$

The fits are undoubtedly better with the modified SIRD model. In particular, while the initial starting phase of the outbreak is captured correctly by both models, the rapid decrease of the infected population at the peak seems to be better captured by enforcing a very quick (about 3 days on average) decrease of the infectivity rate starting from day 40 after 22/01/2020. Moreover, the fit converges to a very small value of  $\alpha$ , suggesting that the containment measures implemented in South Korea have been quite effective.

It is interesting to observe that at the end of February South Korea was apparently under the illusion that the outbreak had levelled off and the number of actual infected was starting to decline (see log-log plots in Fig. 1). It is apparent that the outbreak suddenly reinvigorated, supposedly due to a hotbed that had gone unnoticed. Alternatively, that could have been the date that marked the gear shift to the powerful screening action implemented by the Korean government.

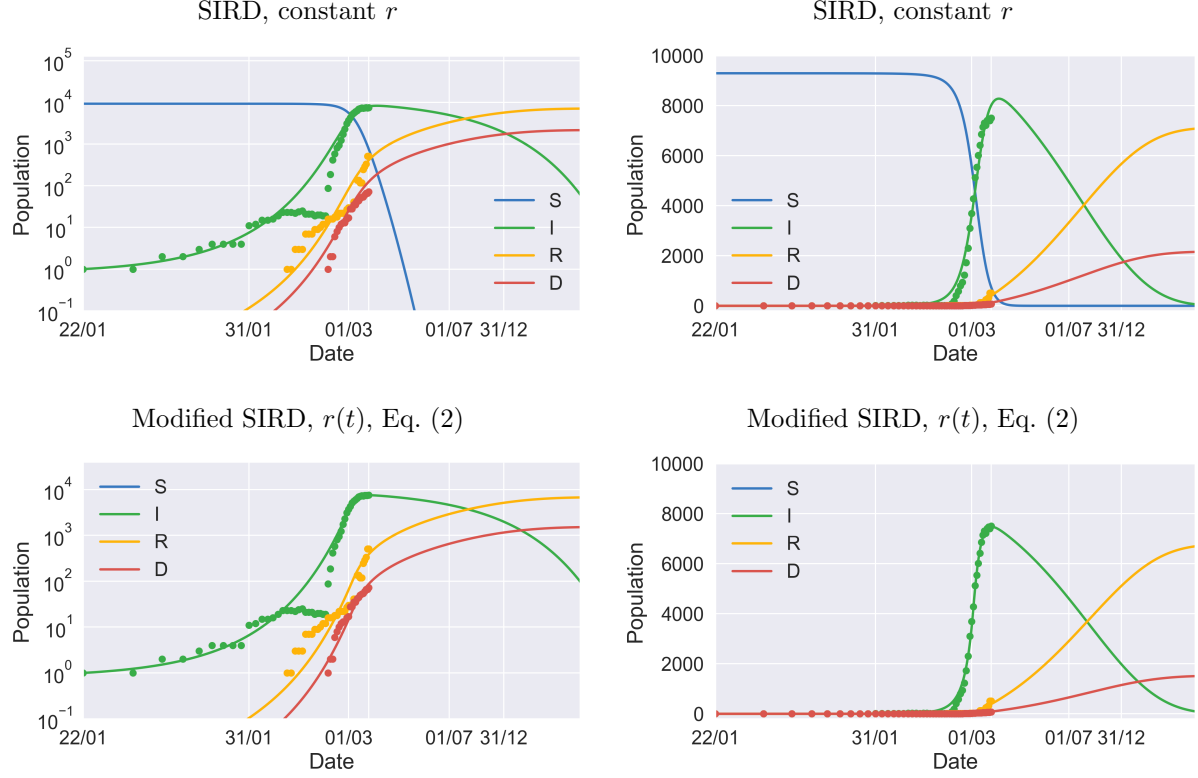


FIG. 1. Predicted evolution of the COVID-19 outbreak in South Korea. Symbols represent the official data retrieved from the CSSE repository [1]. Solid lines are the predicted trends based on the fits of the SIRD model, Eqs (1), to the data (top) and based on the modified SIRD model, Eq. (2) (bottom).

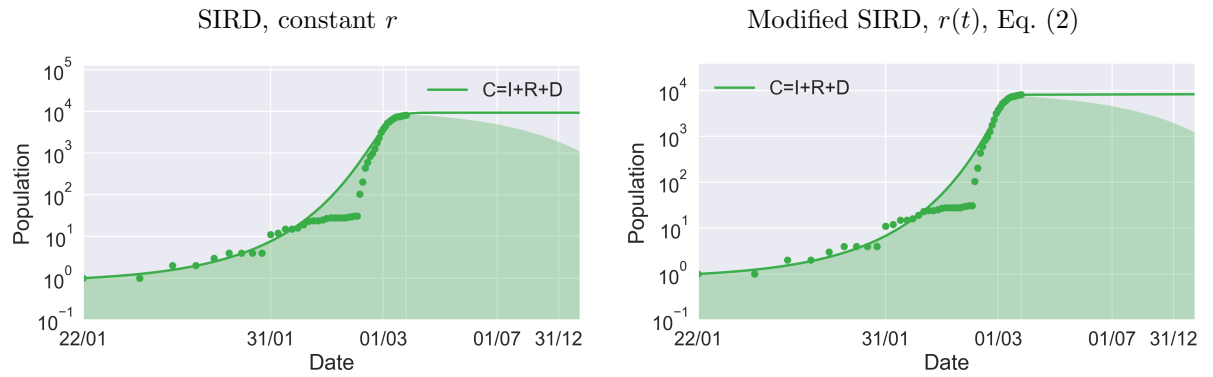


FIG. 2. Predicted evolution of the total number of confirmed infected people for the COVID-19 outbreak in South Korea (solid line). The fitted data are shown as filled circles (see also Table I). The epidemic peak (population  $I$ ) is shown for comparison (transparent filled curve).

Country	$r$ [days <sup>-1</sup> ]	$a$ [days <sup>-1</sup> ]	$d$ [days <sup>-1</sup> ]	$S_0$
<b>Italy</b>	$7.90 \times 10^{-6} \pm 3 \times 10^{-8}$	$2.13 \times 10^{-2} \pm 2 \times 10^{-4}$	$1.63 \times 10^{-2} \pm 2 \times 10^{-4}$	$4.13 \times 10^4 \pm 2 \times 10^2$
<b>China</b>	$3.95 \times 10^{-6} \pm 4 \times 10^{-8}$	$3.53 \times 10^{-2} \pm 1 \times 10^{-4}$	$3.1 \times 10^{-3} \pm 2 \times 10^{-4}$	$8.33 \times 10^4 \pm 2 \times 10^2$
<b>China*</b>	$3.33 \times 10^{-6} \pm 2 \times 10^{-8}$	$1.80 \times 10^{-2} \pm 2 \times 10^{-4}$	$3.0 \times 10^{-3} \pm 2 \times 10^{-4}$	$7.92 \times 10^4 \pm 4 \times 10^2$
<b>S. Korea</b>	$2.458 \times 10^{-5} \pm 8 \times 10^{-8}$	$4.10 \times 10^{-3} \pm 1 \times 10^{-4}$	$1.2 \times 10^{-3} \pm 1 \times 10^{-4}$	$9.310 \times 10^3 \pm 30$
<b>S. Korea<sup>†</sup></b>	$1.80 \times 10^{-6} \pm 7 \times 10^{-7}$	$3.81 \times 10^{-3} \pm 8 \times 10^{-5}$	$7.90 \times 10^{-4} \pm 7 \times 10^{-5}$	$1.40 \times 10^5 \pm 7 \times 10^4$
<b>Lombardy</b>	$1.1029 \times 10^{-5}$	$4.34 \times 10^{-2}$	$2.62 \times 10^{-2}$	$2.6244 \times 10^4$

TABLE I. Table of average values of the best-fit parameters and associated standard deviations computed from 30 independent runs of the stochastic differential evolution algorithm [2], as implemented in the `Python-Scipy` package. The line marked with an asterisk for China refers to a fit limited to the data up to February 19<sup>th</sup> 2020. The data marked by a <sup>†</sup> refer to a fit with the modified SIRD model, Eq. (2). The best-fit values of the parameters for the fit of the outbreak in Lombardy are computed through a deterministic fit algorithm implemented in Matlab (function `fminsearch`).

### III. UPDATED FITTING FOR THE SPREADING OF THE EPIDEMICS IN ITALY.

In the following we report the updated version of the plots included in the main manuscript and relative to the spreading of the epidemics in Italy. As compared to our early predictions, the number of infected individuals at peak grew significantly (as well as the number of susceptible  $S_0$ ). The position of the peak remains however virtually unchanged. This prediction seems indeed rather robust. As an additional complement we report about the situation in the Italian region of Lombardy: the same data are displayed in the two panels of figure 4, respectively in log-lin and in lin-log scale. By inspection, one can appreciate a general trend to deviate from the initial exponential profile. The fitting obtained for the Lombardy region includes data up to the 14/03/2020.

### ACKNOWLEDGMENTS

We are indebted to the many colleagues who keep on sending us insightful observations on this project.

- 
- [1] E. Dong, H. Du, and L. Gardner, *The Lancet Infectious Diseases* **3099**, 19 (2020).
  - [2] R. Storn and K. Price, *Journal of Global Optimization* **11**, 341 (1997).

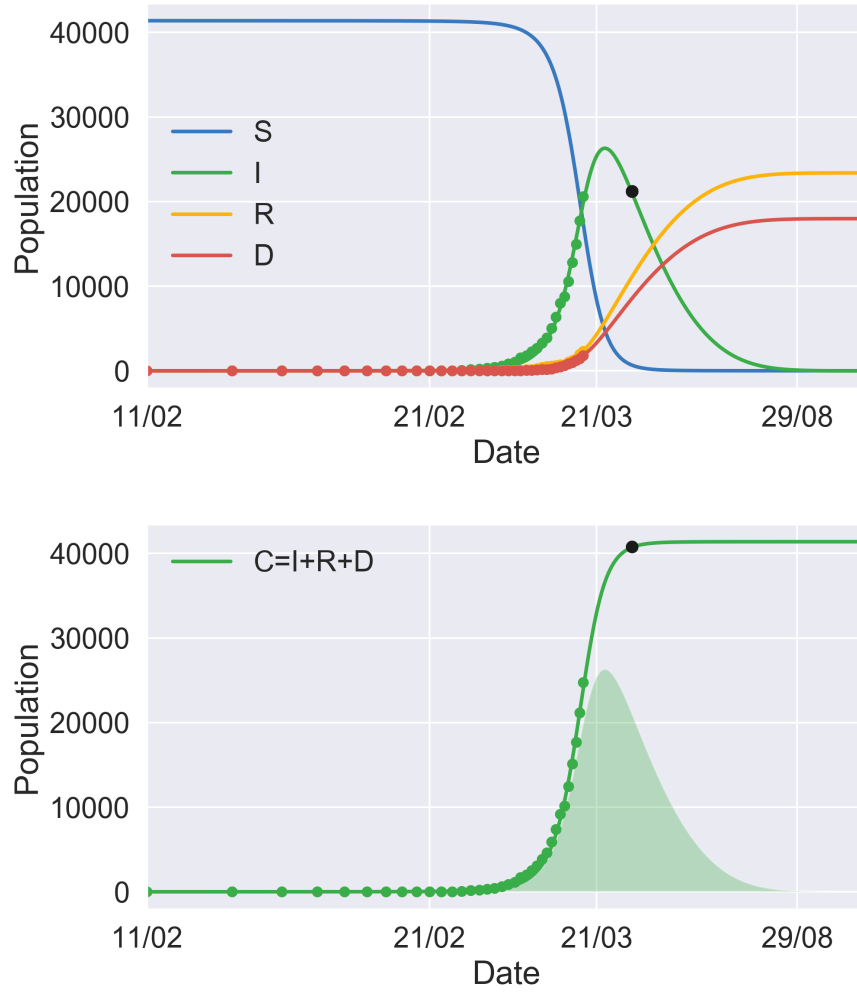


FIG. 3. Predicted evolution of the COVID-19 outbreak in Italy as for the latest available data (15/03/2020). Symbols represent the official data retrieved from the CSSE repository [1]. Solid lines are the predicted trends based on the fits of the SIRD model, Eqs (1), to the data. The black circle in the top graph marks the predicted number of confirmed infected individuals at the announced end of the imposed lockdown on the Italian territory, April 3<sup>rd</sup> 2020. The bottom panel shows the predicted evolution of the *total* number of confirmed infected people  $C = I + R + D$  (solid line), as compared to the data (symbols).

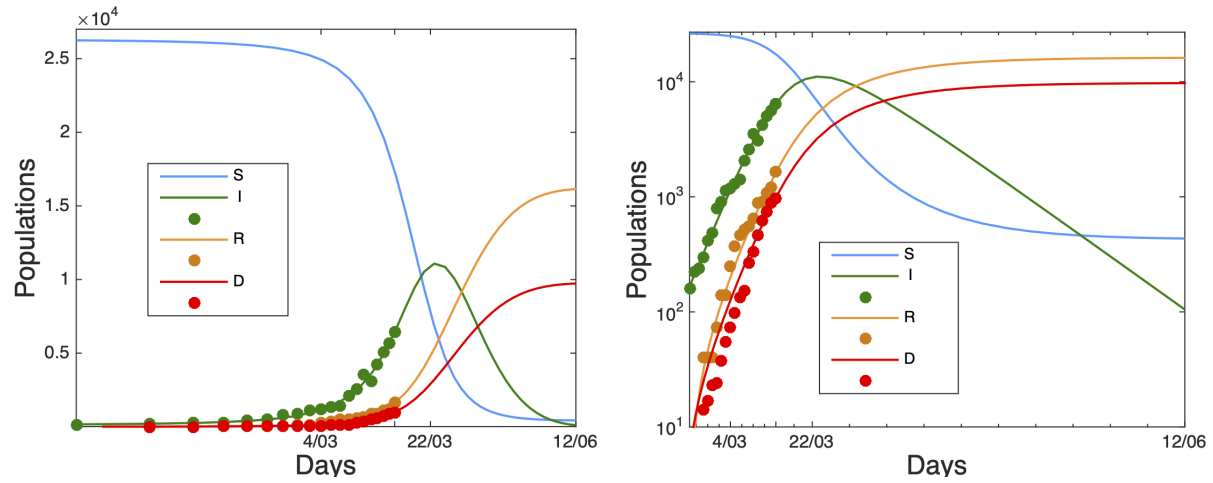


FIG. 4. Predicted evolution of the COVID-19 outbreak in the Italian region of Lombardy. Symbols represent the official data retrieved from the Italian Protezione Civile at: <https://github.com/pcm-dpc/COVID-19/blob/master/dati-regioni/dpc-covid19-ita-regioni.csv>. Solid lines are the predicted trends based on the fits of the SIRD model, Eqs (1), to the data.