

disCOVIDer19

A path-guide inside the COVID-19 pandemia

Fabio Caironi, Marzio De Corato,
Andrea Ierardi, Federico Matteucci,
Gregorio Saporito

April 4, 2020

"The laws of history are as absolute as the laws of physics, and if the probabilities of error are greater, it is only because history does not deal with as many humans as physics does atoms, so that individual variations count for more" (I. Asimov, Foundation and Empire)

CONTENTS

1	Introduction	5
1.1	Background	5
1.2	Historical Background	6
2	Theoretical Background	11
2.1	Population growth dynamics	11
2.2	Application to the epidemiology	12
2.3	Time series approach	13
2.4	An ab-initio approach: the SEIR model	13
3	Data Origin	13
4	Panels	14
4.1	Home	14
4.2	Data Inspection	15
4.3	Data Analysis	20
5	Conclusions	24
6	Packages used	24

LIST OF FIGURES

Figure 1	Different pictures of the virus as reported by the NIAID's Rocky Mountain Laboratories (RML) in Hamilton, Montana [1]	5
Figure 2	Pneumonia and Influenza mortality from 1910-1922 in London: note the three peaks due to the spanish flu. The variation of the diffusion rate R_0 is reported in the box. Image source: Ref. [2]	7
Figure 3	Mortality range distribution for the two waves of Spanish flu: note that the second one was more severe for the young people. Image source: Ref. [3]	8
Figure 4	Wang-ShickRyu diagram for the explanation of the Spanish, Asian flu and Honk-Kong flu . Image source: Ref. [4]	9
Figure 5	Excess mortality due to respiratory disease from 1950 to 1970 as reported by Viboud et al. The dashed lines marks the timing of the Hong Kong flue. Image source: Ref. [5]	10
Figure 6	The Malthusian growth (red line) compared to the logistic growth (blue line). The carrying capacity is marked with a black dashed line. Image taken from [6]	13

Figure 7	Choropleth map by province showing absolute number of Covid-19 cases	14
Figure 8	Input selector where the user can choose the three different breakouts absolute, percentage, and density	15
Figure 9	Boxes with syncronised barometer level statistic about the Coronavirus	15
Figure 10	•	18
Figure 11	19
Figure 12	The age distribution for the recovered people . .	20
Figure 13	Subpanels of Data Analysis panel	22
Figure 14	Subpanels of Data Analysis panel	23

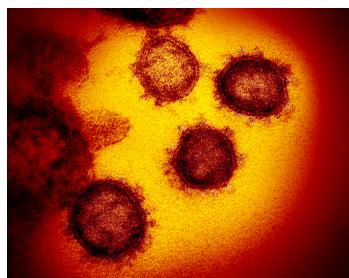
LIST OF TABLES

ABSTRACT

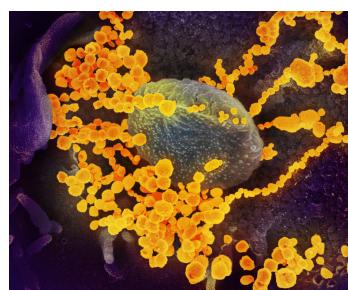
1 INTRODUCTION

1.1 Background

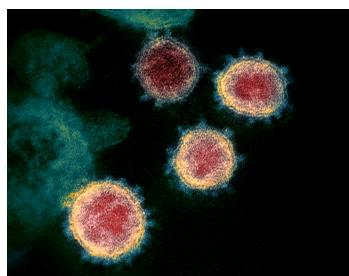
In December 2019 different cases of pneumonia were reported in Wuhan (China) [7]. Their origin was later ascribed to a new virus classified as *Severe acute respiratory syndrome coronavirus 2* (SARS-CoV-2) whose TEM and SEM pictures are reported in the Fig. 1. The origin of this virus is still subject of scientific debate between the scientific community, however one of the most common opinion is that this virus comes from bats, in particular the genus *Rhinolophus* [8]. The most compelling feature of this virus is its ability to spread via coughing and sneezing [9], and also by touching infected surfaces [10]. Differently with respect to the SARS-CoV the virus seems to have a lower mortality rate [11, 12].



(a) Transmission electron microscope (TEM) image of SARS-CoV-2—also known as 2019-nCoV, emerging from the surface of cells cultured



(b) Scanning electron microscope of the SARS-CoV-2 emerging from the surface of cells cultured



(c) TEM image of SARS-CoV-2. Note the spikes that give the name coronavirus to the virus



(d) SEM image of the virus

Figure 1: Different pictures of the virus as reported by the NIAID's Rocky Mountain Laboratories (RML) in Hamilton, Montana [1]

In January the Chinese government imposed the quarantine for the city of Wuhan (almost 11M people); the quarantine was later expanded to the full province of Hubei (60M people) and then to the neighbour provinces Huanggang, Ezhou and Xianning. The virus

then spread in Canada, Germany, Thailand and Japan and then in other different countries including Italy [13]. In Italy the first cases, two Chinese tourist from Hubei, were reported in Rome [14]; then other cases were reported in Codogno (Lodi, Lombardy). Later the virus spread almost in all regions of Italy with a higher density in Lombardy, Veneto, Emilia-Romagna, Piemonte and Marche. Starting from 22nd of February, the Italian government started to impose the quarantine (red-zone) for eleven different municipalities, in particular Codogno, Casalpusterlengo, Lodi (included the neighbour municipalities) and Vo'. From this date onward, different restrictive measures were imposed starting from regions with the highest number of cases: public entrainment were almost suspended as well as schools and universities. Workers (in public and private sectors) were allowed, when possibile, to work at home (smart-working). Such measures culminated with a decree approved by the Italian government that divided Italy in three areas: the red zones, in which the municipalities with highest number of cases in which all the population was subjected to quarantine; the yellow zones (Lombardia, Veneto, Emilia Romagna) where schools, universities and public events, sports as well cinemas were suspended; the rest of Italy, in which no restrictive measures were adopted [15]; however 3 days later all schools and universities were suspended [16]. The restrictive measures were also extended up to all Italy with a further decree [17].

1.2 Historical Background

In the last centuries several cases of pandemic diseases were recorded: from our point of view, their dynamic is useful to understand and perhaps to model that further peaks may succeed the first one even with a highest death rate.

THE DEADLIEST PANDEMICS IN HUMAN HISTORY: THE SPANISH FLU

The first one, which happens to be the biggest human pandemic (in terms of spread just as much as deaths [18]), is the Spanish Flu: this name is quite misleading since it was not originated in Spain, though the Spanish press was the first one to talk about it. Indeed, at that time (1917-1918) almost every other country was involved in the Great War and consequently the censorship was applied to the press. The origin of this disease is still under research: some scholars focused on the United States [19, 20], on France [21] and on Asia [22]. The disease was ascribed to the virus A/H1N1, a subtype of influenza A Virus. The spread of the disease was strongly amplified by the fact that many people lived in very bad hygienic conditions. As one can see from Fig.2 there were basically three waves: the initial spread, the second wave and the third wave. The second wave, that was the most

lethal one, coincides with the end of the war and with the coming back of the troops: the close contact inside the trains, and the diffusion of the troops inside their home villages/cities increased abruptly the contagion rate and then the death rate. Furthermore, this effect was amplified by the fact that a more deadly mutation of the virus became widespread [23]. As pointed out by different scholars, such effect was enhanced by the fact that sick people, which infection potential was higher, were transferred by train to the hospitals. Immediately, the virus spread with a higher rate [24]. As it is possible to note from Fig. 3 this second wave was largely more deadly for young people with respect the first wave. Finally, looking at plot 2, a third wave also manifested itself in the beginning of 1919: here, the mortality was higher with respect to the first one, but lower to the second one. In the end, this pandemic infected 500 million people and killed something between 17 and 50 million [25]. It is worth noting that, just to compare, the casualties in the Great War were from 8.8 to 10.7 million among soldiers and 11 million among civilian [26]

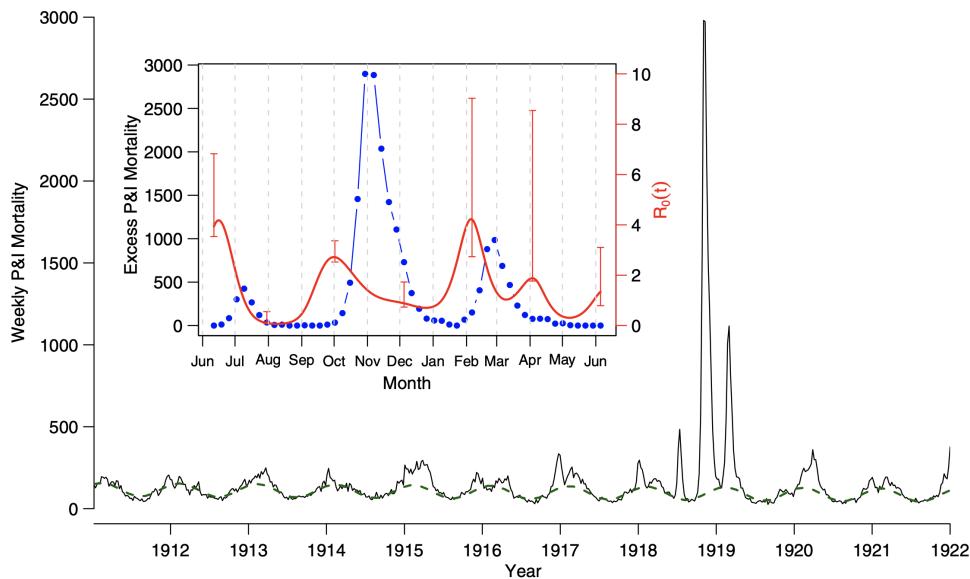


Figure 2: Pneumonia and Influenza mortality from 1910-1922 in London: note the three peaks due to the spanish flu. The variation of the diffusion rate R_0 is reported in the box. Image source: Ref. [2]

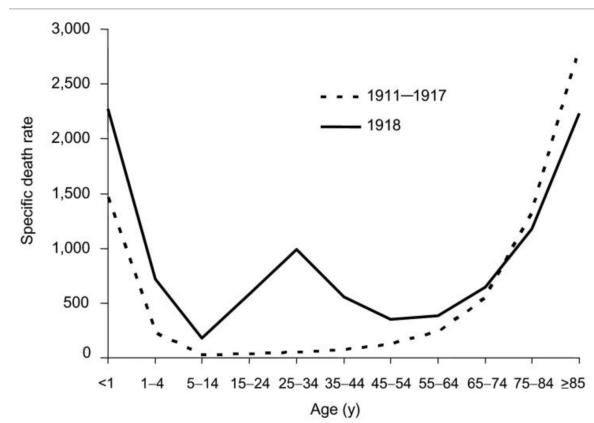


Figure 3: Mortality range distribution for the two waves of Spanish flu: note that the second one was more severe for the young people. Image source: Ref. [3]

THE ASIAN INFLUENZA (H2N2) In 1957 different cases another virus of influenza (A/H₂N₂) spread in China starting from Guizhou [27] . In this case, at difference with respect to the Spanish flu, the scholars were able to isolate the virus (the virus of the former was only isolated in 1933). It was proposed that the virus may be originated from mixing of avian and human influenza (see Fig. 4). The virus diffused in Hong Kong in April then in Singapore, Taiwan and Japan and then globally [28] . While for the US a crucial role for the spread was supposed to be played by naval bases [29, 28], for Europe the same role was played by the land route through Asia [30, 28]. Scholar also pointed out the role of a church conference in Iowa [31] where almost 1600 people from all countries met. Furthermore the opening of school in fall increased abruptly the spread of the virus (40-60 % clinical attack rate) [32, 28]. Also in this case it was reported a second wave (or perhaps others one, see Fig. 5) more lethal than the first one [33] . In this case, contrarily to the Spanish Flu a vaccine was developed although [32], it was argued that its diffusion was not effective to diminish the effect of the pandemics [34]. It was nothing that according to scholars [35] no effective measures of quarantine or other non-pharmacological one were undertaken by the authorities [35] due to the mild symptoms [28]. The total toll for this pandemic was from 1 to 1.4 million deaths according to the WHO [36]: from our point of view this pandemic is a counterfactual teaching lesson about what happens when a pandemic is underestimated (and so no restrictive measure is taken) due to its relative mild symptoms

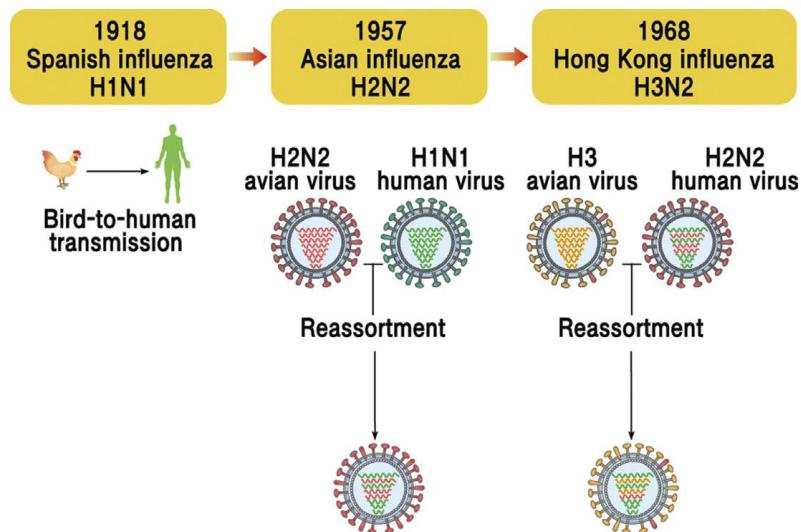


Figure 4: Wang-ShickRyu diagram for the explanation of the Spanish, Asian flu and Honk-Kong flu . Image source: Ref. [4]

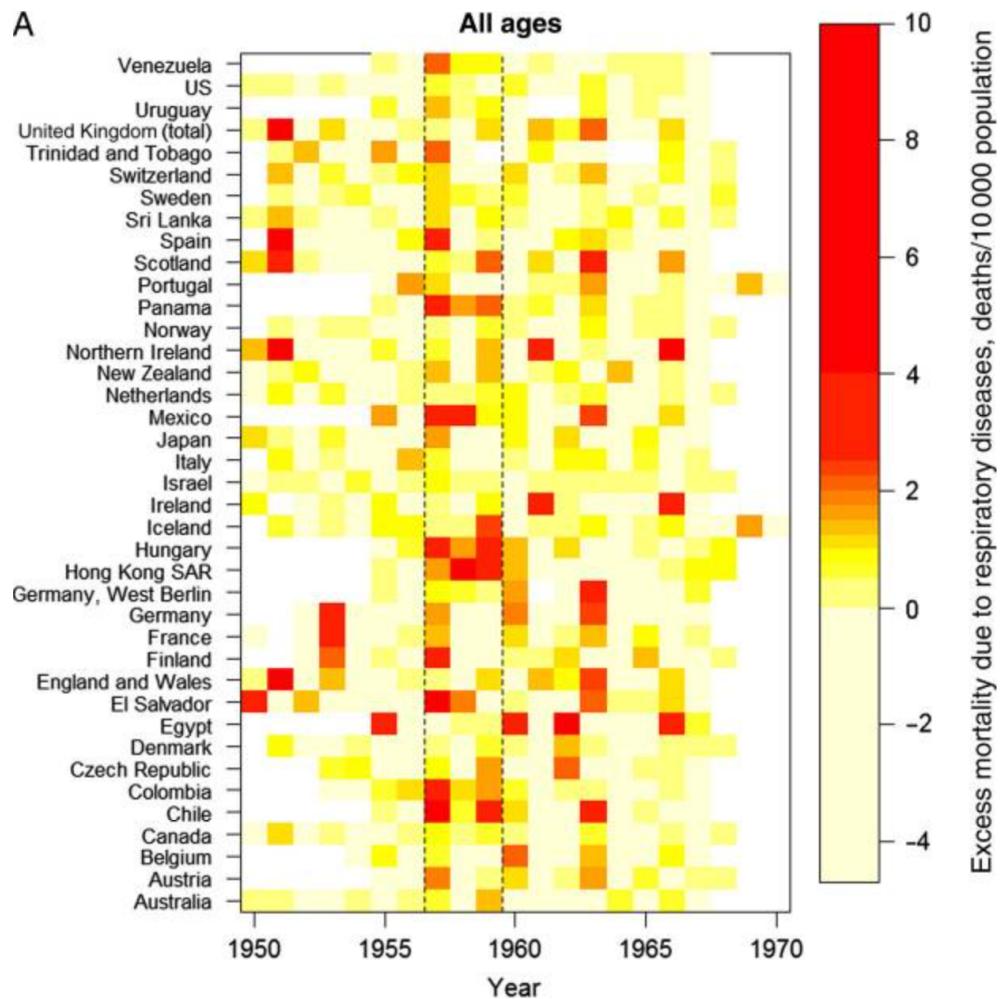


Figure 5: Excess mortality due to respiratory disease from 1950 to 1970 as reported by Viboud et al. The dashed lines marks the timing of the Hong Kong flue. Image source: Ref. [5]

MOTIVATION In order to have a thorough insight of the spread of the COVID-19 disease and its effects on Italian population, we propose a ShinyApp, called disCOVIDer19, that gathers, presents and analyzes the numbers of new and cumulative infections, hospitalized and intensive care patients, deaths and recoveries, over the whole country and over its individual regions and provinces. A detailed guide to the functioning of the App is available in sections 3 and 4. As regards the analysis panel, we fitted the cumulative cases with a logistic curve, in order to get a bird-of-eye view on the general and specific trends as well as on the effectiveness of the restrictive measures imposed by the Italian government. The theoretical foundations of this fit and how to manage and use it will be discussed in the section 2 and 4.3. Beside the logistic curve, we also considered a statistical approach, largely diffused among economists, that considers the number of total cases in each day as a time series: on this basis we were able to make use of a particular tool, which will be briefly introduced in the section 2, that allowed us to make a further forecast about future cases. This latter becomes useful in case of large deviations from logistic distribution. Finally, we must remark that our approach is informational: we do limit ourselves to list and comment our findings without claiming rather delicate conclusions, especially about predictions, which many people hinge on nowadays and many lives too, unfortunately.

2 THEORETICAL BACKGROUND

Several models were proposed in the literature for modelling the spread of a disease [37]: here, we are going to consider a relatively simple model taken from the growth dynamics of populations.

2.1 Population growth dynamics

The dynamics of population was founded by Thomas Malthus in 1817 [38] with his well known equation:

$$x_{n+1} = x_n(1+r) \quad (1)$$

where x_n stands for the population at time and r the growth rate. Renaming x to P and applying recursive substitution, we get:

$$P_{n+1} = P_0(1+r)^{n+1} \quad (2)$$

in which P_0 represents the initial population. This discrete model can be reshaped into a continuous one in the following way:

$$\dot{P}(t) = rP(t) \quad (3)$$

That's a first order ODE with separable variable. Integrating yields:

$$P(t) = P_0 e^{rt} \quad (4)$$

Malthus, of course, did not believe that the population could grow ad infinitum with an exponential trend and, since he estimated that the resources growth follows a linear path, he argued that, at the intersection of the two curves, a consistent part of the population will not have access to the resources. Thus he expected that the exponential growth realizes only in the first part of the growth. Basing on these arguments, in 1838 Belgian mathematician Pierre-François Verhulst developed a model [39] whose new feature and key point was the maximum size of the population allowed by resources, that is the natural upper limit of coexisting individuals in a particular environment, usually called the *carrying capacity*, here indicated as K (P and r have the same meaning of the previous model):

$$\dot{P} = rP \left(1 - \frac{P}{K}\right) \quad (5)$$

that has the following analytic solution:

$$P = \frac{K P_0 e^{rt}}{K + P_0 (e^{rt} - 1)} \quad (6)$$

In Fig.6 this solution is compared to the Malthusian one: as one can see, the logistic growth shows a saturation effect towards the asymptote of carrying capacity. This completes Malthus's intuition, as it represents population growth before and after the intersection between the linear resources curve and the natural exponential growth. It is worth noting that in a neighbourhood of the origin the two curves are identical; later, the logistic curve bends down, still increasing, till it reaches the saturation region, in which it becomes almost flat.

2.2 Application to the epidemiology

The dynamics of the two models described before can be used to shape the spread of an infection [40, 41] in this case the population P is replaced with the total number of infected people, P_0 with the number of initially infected people, r with the spreading rate and the carrying capacity K with the maximum number of people that can be infected. The latter is the key point for the modelling: in principle, this parameter would be equal to the population number; in practice, due to restrictive measures and, possibly, vaccines, a large share of the total population won't enter the computation, thus lowering the carrying capacity. So, we can argue that the effectiveness of the restrictive measures can be inferred from the behaviour of the cumulative curve (except, of course, if the number of infected people is close to

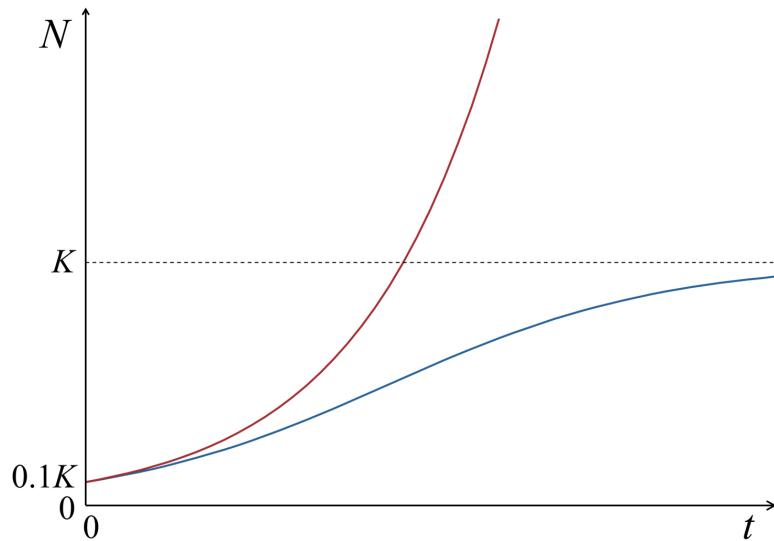


Figure 6: The Malthusian growth (red line) compared to the logistic growth (blue line). The carrying capacity is marked with a black dashed line. Image taken from [6]

the population size). It is worth noting that, for each non-quarantined infected, the carrying capacity is increased up according to the individuals that can be infected by these subjects. If this contribution is not negligible a new logistic growth may be found, with of course an exponential starting behaviour. Thus the prediction with the logistic curve should be taken as an evaluation of the impact of restrictive measures and as the *best scenario* that may occur. For this reason it is advisable to see the logistic curve as a local (in time) estimator. In fact, one can model a differential set of equation were the carrying capacity is time dependent. This possibility may be considered by the authors in a future version of this App. Furthermore, we remind that other similar models such as Gopertz, Richards or Bertalanffy can be used to shape the COVID-19 outbreak[41]: the authors may consider to include them in a future version of this App.

2.3 Time series approach

2.4 An ab-initio approach: the SEIR model

3 DATA ORIGIN

4 PANELS

4.1 Home

The home panel provides an overview of the COVID19 spread in Italy. The data shown are synchronised with the civil protection database. Whenever the App is started, a check for the updates is performed and their occurrence is indicated in the *most recent updates* section. The home panel consists of two main sections: the choropleth map and the summary statistics. The choropleth map (Fig. 7) is an interactive heat map with breakouts by region and by province tracking the number of Covid-19 cases.

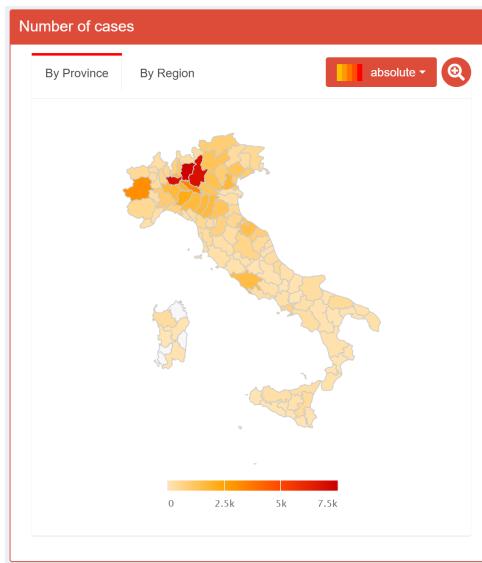


Figure 7: Choropleth map by province showing absolute number of Covid-19 cases

Beyond the raw absolute number of total cases by province/region we propose two other indicators which introduce some form of normalisation to improve the comparability of the different geographical areas. These indicators are percentage and density (Fig. 8).

The percentage indicator is calculated dividing the total number of COVID19 cases by the population of the respective region/province (the latter retrieved from Istat [42, 43]).

$$\text{percentage cases} = \frac{\text{total cases}}{\text{local population}} \times 100 \quad (7)$$

This indicator ensures that the number of cases in more populated areas are not over-indexed compared to less populated ones. Similarly, it ensures less populated areas are not under-indexed. The density indicator accounts for the territorial extension (in km^2) of re-

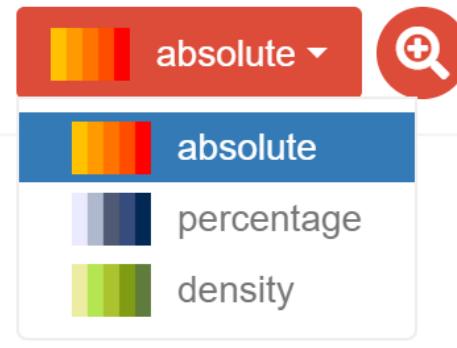


Figure 8: Input selector where the user can choose the three different break-outs absolute, percentage, and density

gions/provinces expressed in parts per thousand (%) and is calculated as follows:

$$\text{density cases} = \frac{\text{total cases}}{\text{territorial extension}} \times 1000 \quad (8)$$

This indicator combines the population density with infection spread, offering a control over space. Indeed, it is directly proportional to population density and to percentage cases:

$$\begin{aligned} \text{density cases} &= \frac{\text{total cases}}{\text{local population}} \cdot \frac{\text{local population}}{\text{territorial extension}} \times 1000 \\ &= \text{percentage cases} \cdot \text{density population} \times 1000 \end{aligned} \quad (9)$$

It is enhanced by densely populated territories, which, by the way, may in turn enhance spread infection too. Finally, the summary statistics section (Fig. 9) provides some barometer level statistics regarding the current total number of COVID19 cases in Italy with three further breakouts: intensive care, hospitalised, and home isolation.



Figure 9: Boxes with synchronised barometer level statistic about the Coronavirus

4.2 Data Inspection

The *Data Inspection* panel provides an exploratory analysis of the COVID19 spread in Italy and it consists of 2 main sections: an introductory inspection followed by a deeper inspection.

The first section introduces the user to a higher-level time series with the daily frequency of the number of Covid-19 cases broken down into symptomatic, recovered, deaths, and intensive care (Fig. 11a in the *General info* box). The cumulative cases add up each day while the daily cases are obtained by taking the first difference of the cumulative series ($y_t - y_{t-1}$). When these time series capture all the stages of the pandemic evolution, the daily data are expected to assume a bell-shaped curve with a peak while the cumulative series a shape that is more similar to a logistic curve. The user can select different options spanning from the macro-level country view to the more granular province-level. In the same *General Info Box*, the second tab provides direct access to the raw data (Fig. 10b) which is behind most of the visualisation used in the *Home* and *Inspection* sections of the dashboard. Through some input selectors and buttons, the dataset can be queried to show the information that is of most interest for the user. The *Infected Age Distribution* box

The next chart shows the age distribution of infected patients (Fig. [?]) with a region selector that offers the possibility to further investigate local-level dynamics and spot regional differences. The data was retrieved from Istituto Superiore di Sanità (ISS) [44] which is the National Institute of health in Italy. The information dates back to the 30th of March 2020. The chart shows the age distribution of the population which is divided into 10 age bins each one representing a 10-year interval. In particular, the bar chart indicates the absolute number of infected people in the respective age category while the pie chart shows the percentage breakdown of cases.

The *Deeper Inspection* section consists of 4 different types of visualisations. The first two are located in the two tabs of the box called *Intensive Care Information*. As the title suggests, these charts provide a comparative visualisation across regions of intensive care occupancy in hospitals. While one chart shows the occupancy/capacity index in percentage terms (Fig. 10a), the other shows absolute figures of capacity and occupancy by region (Fig. 11c). In particular, the occupancy index is calculated by dividing the percentage occupancy in the selected day by the hospital initial intensive care capacity at the start of the pandemic.

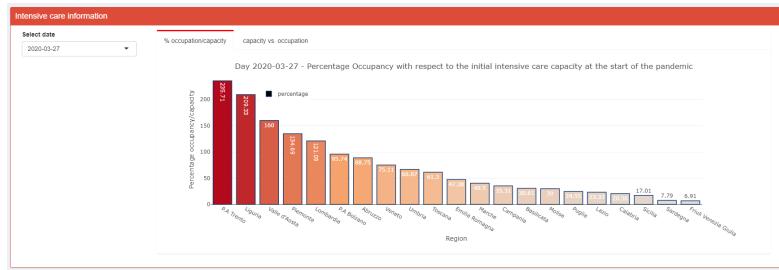
$$\text{percentage occupancy} = \frac{\text{occupancy}}{\text{capacity}} \times 100 \quad (10)$$

It is worth noting that occupancy may be higher than the total number of available spots due to the fact that the beds in intensive care may have been more recently upgraded (see Sec. 3). The date selector, allows the user to explore the time evolution and variations of such

quantities. The third chart (Fig.11b) represents the percentage growth and growth.

$$\text{growth} = \frac{\text{cases}_t - \text{cases}_{t-1}}{\text{cases}_t} \quad \text{growth}\Delta = \text{growth}_t - \text{growth}_{t-1} \quad (11)$$

Despite eventual growth in the number of infected people, a persistent negative Δ is a useful indicator to signal if the pandemic spread is slowing down. The fifth chart (10c) visualises the daily cases with respect to the daily tests. It is useful to discern if an increased number of cases is due to an endogenous increase in the number of tests or due to an exogenous increase in cases due to new outbreaks.



(a)

Plot Raw data

Select dates: 2020-02-24 to 2020-03-27

National By region By province

Lombardia Show

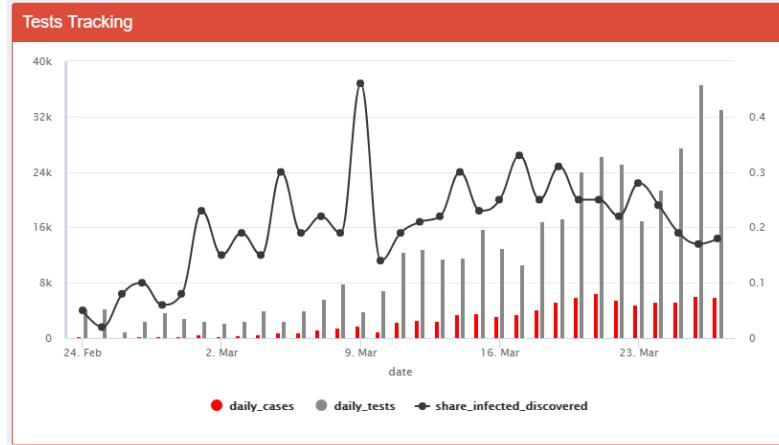
Show 10 entries

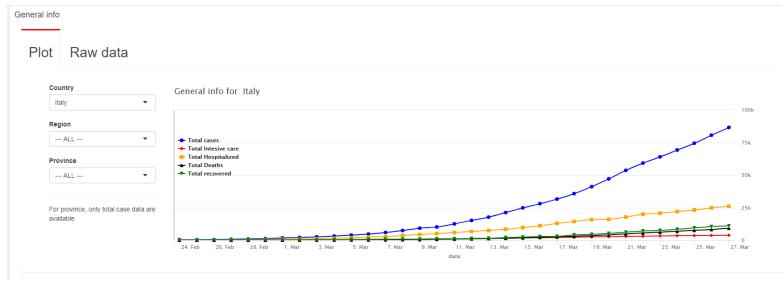
data	ricoverati_con_sintomi	terapia_intensiva	totale_ospedalizzati	isolamento_domiciliare	totale_attualmente_positivi	nuovi_attualmente_positivi	dismessi_guariti	deceduti	totale_casi
1 2020-02-24	76	19	95	71	166	166	0	6	166
2 2020-02-25	79	25	104	127	231	65	0	9	231
3 2020-02-26	79	25	104	145	249	18	0	9	249
4 2020-02-27	172	41	213	136	349	100	40	14	349
5 2020-02-28	235	47	282	192	474	125	40	17	474
6 2020-02-29	256	80	336	216	852	78	40	23	852
7 2020-03-01	406	106	512	375	887	335	73	24	887
8 2020-03-02	478	127	605	472	1077	190	139	38	1077
9 2020-03-03	698	167	865	461	1326	249	139	55	1326
10 2020-03-04	877	209	1085	411	1497	171	230	73	1497

Showing 1 to 10 of 53 entries

Previous 1 2 3 4 Next

(b)

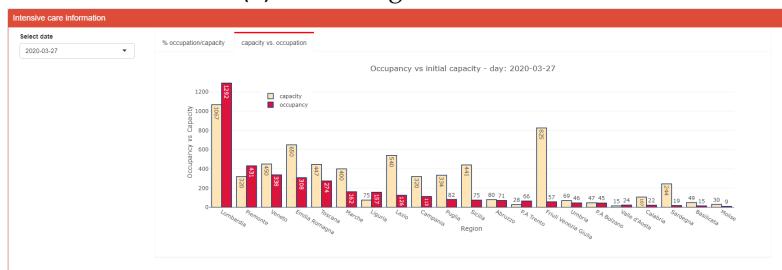




(a) SEM image of the virus



(b) SEM image of the virus



(c) SEM image of the virus

Figure 11

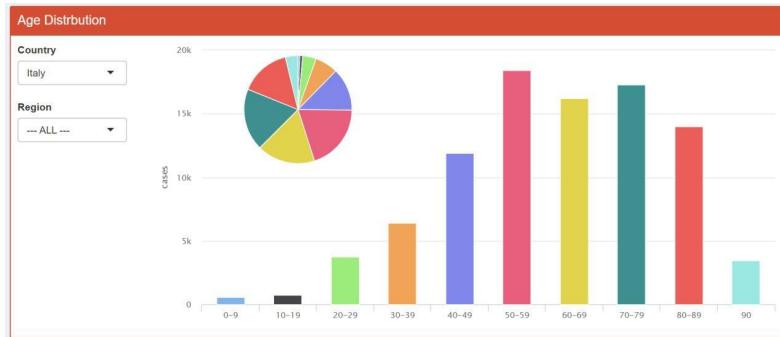


Figure 12: The age distribution for the recovered people

4.3 Data Analysis

Panel *Analysis* explores different mathematical models for shaping the infection's time series, namely the ones introduced in Section 2. As previously argued, each of these model's predictive power hinges on uncertain actual scenarios, such as the effectiveness of restrictive measures, the presence of several independent and spread outbreaks, and so on.

Package [45] has been used for growth curve fitting. This package provides function *SummarizeGrowth*, which inputs the time series of sample data and outputs the model represented by the growth metrics. Such parameters include those of the logistic equation that best fit the data, namely K , P_0 and r , as in Eq.6. The reasons why this package was chosen are chiefly two. At the one hand, it performs non-linear curve fitting with the Levenberg-Marquardt algorithm [46], which happens to be one of the most robust nls algorithms. For this step, function [47] is used. On the other hand, the package is endowed with methods for background correction.

At the top of analysis panel the user is allowed to pick one dataset among national, by region and by province. Hence, any following analysis will use the selected dataset as input (the user can browse the raw dataset in panel inspection, see paragraph 4.2).

The first section is devoted to the logistic model. An input box (Fig.13a) lists several dynamic inputs:

1. A date range that the calculator will use for curve fitting. The option to choose initial and final dates different from those at, respectively, the beginning and the end of the dataset, is useful in two ways:
 - It allows to cut initial dates which possibly correspond to zero or constant infections data, as a consequence of a delay in the outbreak in that territory. Actually, an automatic

process will do so after the choice of territory, adjusting the initial date to the first day whose new infections exceed a given threshold (currently set to 1).

- It allows to compute the available prediction n days ahead, by cutting the last n dates, and compare it to the real data.
2. A check-box *standardise positive cases by total swabs*. If selected, the share of infected among the tested, instead of the infected themselves, will be used for computations and rendering. This may be useful to capture the selection bias, but only in the scenario in which tests are randomly assigned to the population, or at least are assigned with a logic common to all dates and territories. This option is not currently available for provinces.
 3. Plot type selection boxes. The available plot types are cumulative cases and new cases. (The user can always show and hide plot's components by clicking on their label in the legend).
 4. Residuals plot type. Four options for graphically rendering the residuals: Residuals vs. fitted values; Standardized residuals vs. fitted values; Autocorrelation; Square root of absolute residuals vs. fitted values.

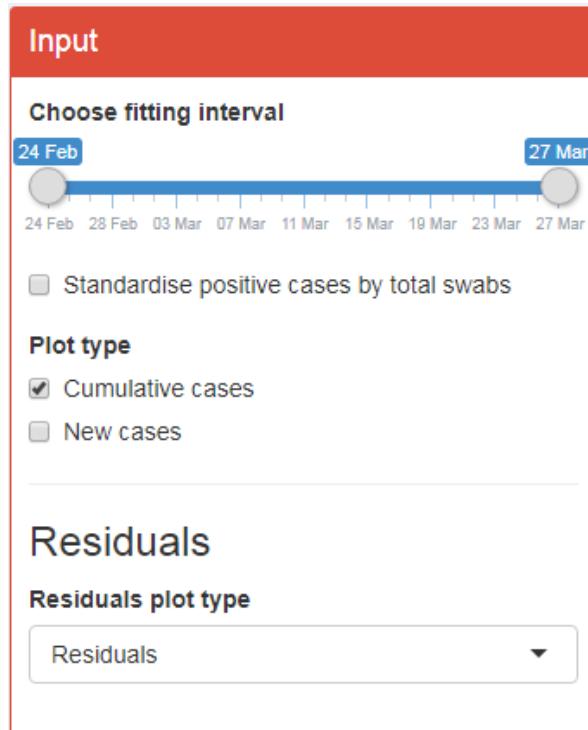
The output is displayed within three boxes:

1. Inside the [13b](#) box the sample points, the fitted logistic curve and a confidence band at 95 % level for the logistic curve are overlayed. The user can view the sample and fitted values by hovering over or touching (for touch-screen devices) them [13b](#). In addition, if *new cases* is selected, a plot of sample cases differences (which correspond to new cases, in fact) is shown over the logistic estimated distribution [13c](#). The latter is simply obtained by deriving the right hand side of formula (4.3.1):

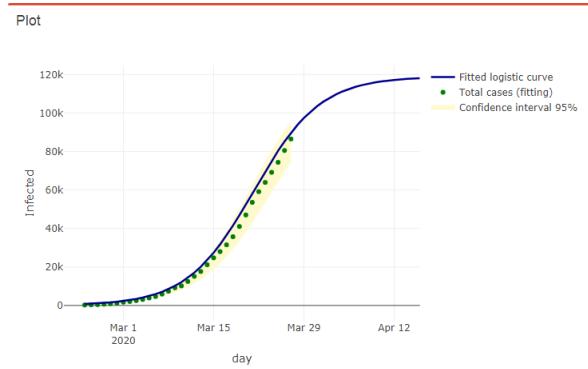
$$N'(t) = \frac{(\frac{K-N_0}{N_0})Kre^{-rt}}{[1 + (\frac{K-N_0}{N_0})e^{-rt}]^2} \quad (12)$$

The user is suggested to take full advantage of plotly features. Besides showing points labels when passing over it and enabling plot selection by clicking on label items, plotly charts are endowed with a sequence of tools shown at the top-right of the box [14a](#).

2. Summary output box [14c](#) contains principal information about the selected model and a list of goodness-of-fit tests on the residuals.
3. Residuals box contains the plots of residuals, rendered in the user-chosen fashion. (see Fig.[14b](#)).



(a) Set of useful tools for navigating into plots, by plotly.

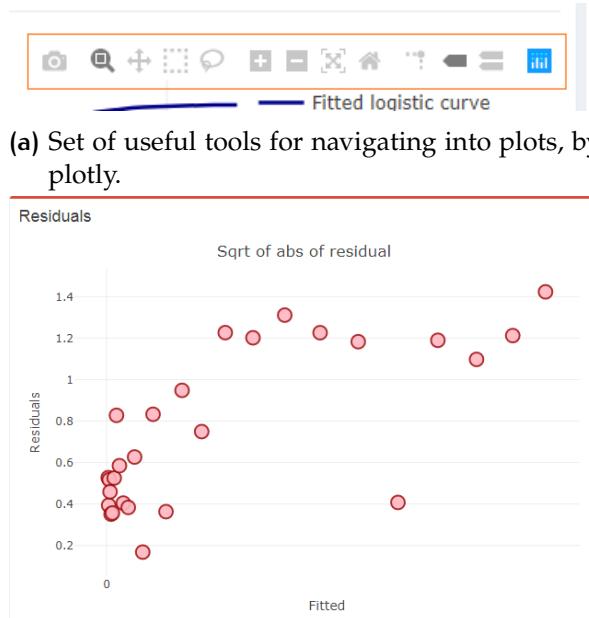


(b) Summary output box of logistic section.

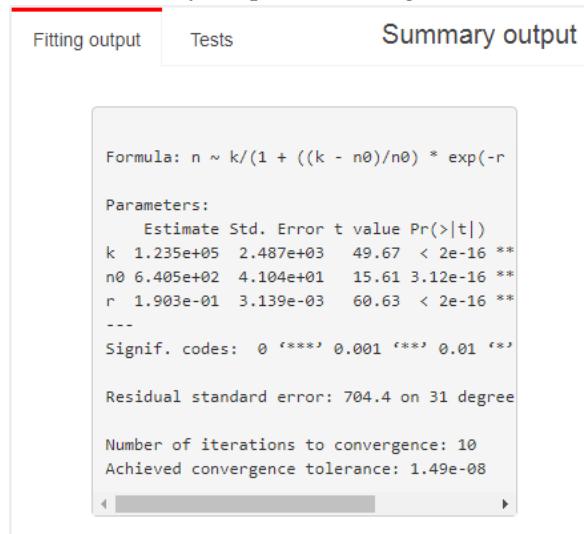


(c) One plot of nls residuals: square root of absolute residuals vs. fitted values.

Figure 13: Subpanels of Data Analysis panel



(b) Summary output box of logistic section.



(c) One plot of nls residuals: square root of absolute residuals vs. fitted values.

Figure 14: Subpanels of Data Analysis panel

5 CONCLUSIONS

6 PACKAGES USED

REFERENCES

- [1] <https://www.niaid.nih.gov/news-events/novel-coronavirus-sarscov2-images>.
- [2] Daihai He, Jonathan Dushoff, Troy Day, Junling Ma, and David JD Earn. Mechanistic modelling of the three waves of the 1918 influenza pandemic. *Theoretical Ecology*, 4(2):283–288, 2011.
- [3] Jeffery K Taubenberger and David M Morens. 1918 influenza: the mother of all pandemics. *Emerging infectious diseases*, 12(1):15, 2006.
- [4] Wang-Shick Ryu. Chapter 15 - influenza viruses. In Wang-Shick Ryu, editor, *Molecular Virology of Human Pathogenic Viruses*, pages 195 – 211. Academic Press, Boston, 2017.
- [5] Cécile Viboud, Lone Simonsen, Rodrigo Fuentes, Jose Flores, Mark A Miller, and Gerardo Chowell. Global mortality impact of the 1957–1959 influenza pandemic. *The Journal of infectious diseases*, 213(5):738–745, 2016.
- [6] https://commons.wikimedia.org/wiki/File:Malthusian_growth_vs_logistic_growth.png.
- [7] Chaolin Huang, Yeming Wang, Xingwang Li, Lili Ren, Jianping Zhao, Yi Hu, Li Zhang, Guohui Fan, Jiuyang Xu, Xiaoying Gu, et al. Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. *The Lancet*, 395(10223):497–506, 2020.
- [8] Peng Zhou, Xing-Lou Yang, Xian-Guang Wang, Ben Hu, Lei Zhang, Wei Zhang, Hao-Rui Si, Yan Zhu, Bei Li, Chao-Lin Huang, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, pages 1–4, 2020.
- [9] Isaac Ghinai, Tristan D McPherson, Jennifer C Hunter, Hannah L Kirking, Demian Christiansen, Kiran Joshi, Rachel Rubin, Shirley Morales-Estrada, Stephanie R Black, Massimo Pacilli, et al. First known person-to-person transmission of severe acute respiratory syndrome coronavirus 2 (sars-cov-2) in the usa. *The Lancet*, 2020.
- [10] De Chang, Huiwen Xu, Andre Rebaza, Lokesh Sharma, and Charles S Dela Cruz. Protecting health-care workers from sub-clinical coronavirus infection. *The Lancet Respiratory Medicine*, 8(3):e13, 2020.

- [11] Morten Dræby Sørensen, Brian Sørensen, REGINA GONZALEZ-DOSAL, Connie Jenning Melchjorsen, Jens Weibel, Jing Wang, Chen Wie Jun, Yang Huanming, and Peter Kristensen. Severe acute respiratory syndrome (sars) development of diagnostics and antivirals. *Annals of the New York Academy of Sciences*, 1067(1):500–505, 2006.
- [12] Paul Weiss and David R Murdoch. Clinical course and mortality risk of severe covid-19. *The Lancet*, 2020.
- [13] <https://hgis.uw.edu/virus/>.
- [14] <https://www.wired.co.uk/article/coronavirus-italy>.
- [15] https://www.repubblica.it/politica/2020/03/01/news/coronavirus_misure_governo-249980561/.
- [16] https://www.repubblica.it/politica/2020/03/04/news/coronavirus_governo_chiuse_le_scuole_in_tutta_italia_-250209485/.
- [17] http://www.salute.gov.it/portale/news/p3_2_1_1_1.jsp?lingua=italiano&menu=notizie&p=dalministero&id=4186.
- [18] Christopher W Potter. A history of influenza. *Journal of applied microbiology*, 91(4):572–579, 2001.
- [19] Alfred W. Crosby. *America's Forgotten Pandemic: The Influenza of 1918*. Cambridge University Press, 2 edition, 2003.
- [20] John M Barry. The site of origin of the 1918 influenza pandemic and its public health implications. *Journal of Translational medicine*, 2(1):3, 2004.
- [21] G Dennis Shanks. No evidence of 1918 influenza pandemic origin in chinese laborers/soldiers in france. *Journal of the Chinese Medical Association*, 79(1):46–48, 2016.
- [22] Christopher Langford. Did the 1918–19 influenza pandemic originate in china? *Population and Development Review*, 31(3):473–505, 2005.
- [23] John M Barry, Cécile Viboud, and Lone Simonsen. Cross-protection between successive waves of the 1918–1919 influenza pandemic: epidemiological evidence from us army camps and from britain. *The Journal of infectious diseases*, 198(10):1427–1434, 2008.

- [24] Peter C Wever and Leo Van Bergen. Death from 1918 pandemic influenza during the first world war: a perspective from personal and anecdotal evidence. *Influenza and other respiratory viruses*, 8(5):538–546, 2014.
- [25] Peter Spreeuwenberg, Madelon Kroneman, and John Paget. Re-assessing the global mortality burden of the 1918 influenza pandemic. *American journal of epidemiology*, 187(12):2561–2567, 2018.
- [26] Philip J Haythornthwaite. *The World War One Source Book*. Arms and Armour, 1996.
- [27] Lachlan M Strahan. An oriental scourge: Australia and the asian flu epidemic of 1957. *Australian Historical Studies*, 26(103):182–201, 1994.
- [28] Patrick R Saunders-Hastings and Daniel Krewski. Reviewing the history of pandemic influenza: understanding patterns of emergence and transmission. *Pathogens*, 5(4):66, 2016.
- [29] DA Henderson. The development of surveillance systems. *American journal of epidemiology*, 183(5):381–386, 2016.
- [30] Alexander D Langmuir. Epidemiology of asian influenza: with special emphasis on the united states. *American Review of Respiratory Disease*, 83(2P2):2–14, 1961.
- [31] AM Payne. Some aspects of the epidemiology of the 1957 influenza pandemic., 1958.
- [32] Donald A Henderson, Brooke Courtney, Thomas V Inglesby, Eric Toner, and Jennifer B Nuzzo. Public health and medical responses to the 1957–58 influenza pandemic. *Biosecurity and bioterrorism: biodefense strategy, practice, and science*, 7(3):265–273, 2009.
- [33] <http://www.britannica.com/event/Asian-flu-of-1957>.
- [34] KE Jensen, FL Dunn, and RQ Robinson. Influenza, 1957: a variant and the pandemic. *Progress in medical virology. Fortschritte der medizinischen Virusforschung. Progres en virologie medicale*, 1:165–209, 1958.
- [35] YATES TROTTER Jr, Frederick L Dunn, Robert H Drachman, Donald A Henderson, Mario Pizzi, Alexander D Langmuir, et al. Asian influenza in the united states, 1957–1958. *American journal of hygiene*, 70(1):34–50, 1959.
- [36] https://apps.who.int/gb/ebwha/pdf_files/WHA64/A64_10-en.pdf.

- [37] Matt J Keeling and Pejman Rohani. *Modeling infectious diseases in humans and animals*. Princeton University Press, 2011.
- [38] Thomas Robert Malthus. *An essay on the principle of population, as it affects the future improvement of society. With remarks on the speculations of mr. Godwin, m. Condorcet, and other writers. By TR Malthus.* 1817.
- [39] PF Verhulst. Notice sur la loi que la population poursuit dans son accroissement in: Correspondance mathématique et physique, vol. 10. 1838.
- [40] Robert E Serfling. Historical review of epidemic theory. *Human biology*, 24(3):145–166, 1952.
- [41] Junling Ma, Jonathan Dushoff, Benjamin M Bolker, and David JD Earn. Estimating initial epidemic growth rates. *Bulletin of mathematical biology*, 76(1):245–260, 2014.
- [42] [https://www.istat.it/it/files//2015/04/
Superfici-delle-unit%C3%A0-amministrative-Dati-comunali-e-provinciali.
zip.](https://www.istat.it/it/files//2015/04/Superfici-delle-unit%C3%A0-amministrative-Dati-comunali-e-provinciali.zip)
- [43] [http://dati.istat.it/Index.aspx?DataSetCode=DCIS_
POPRES1.](http://dati.istat.it/Index.aspx?DataSetCode=DCIS_POPRES1)
- [44] [https://www.epicentro.iss.it/coronavirus/bollettino/
Bollettino-sorveglianza-integrata-COVID-19_
30-marzo-2020_appendix.pdf.](https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19_30-marzo-2020_appendix.pdf)
- [45] [https://cran.r-project.org/web/packages/growthcurver/
index.html.](https://cran.r-project.org/web/packages/growthcurver/index.html)
- [46] Jorge J Moré. The levenberg-marquardt algorithm: implementation and theory. In *Numerical analysis*, pages 105–116. Springer, 1978.
- [47] [https://www.rdocumentation.org/packages/minpack.lm/
versions/1.2-1/topics/nlsLM.](https://www.rdocumentation.org/packages/minpack.lm/versions/1.2-1/topics/nlsLM)