On the uncertainty of real-time predictions of epidemic growths: a COVID-19 case study for China and Italy

Tommaso Alberti[a], Davide Faranda[b,c,d]

[a]*INAF - Istituto di Astrofisica e Planetologia Spaziali, via del Fosso del Cavaliere 100, 00133 Roma, Italy*
[b]*Laboratoire des Sciences du Climat et de l'Environnement, 5 CEA Saclay l'Orme des Merisiers, UMR 8212 CEA-CNRS-UVSQ, 6 Université Paris-Saclay & IPSL, 91191 Gif-sur-Yvette, France*
[c]*London Mathematical Laboratory, 8 Margravine Gardens, London, W6 8RH, UK*
[d]*LMD/IPSL, Ecole Normale Superieure, 9 PSL research University, Paris, France*

Abstract

While COVID-19 is rapidly propagating around the globe, the need for providing real-time forecasts of the epidemics pushes fits of dynamical and statistical models to available data beyond their capabilities. Here we focus on statistical predictions of COVID-19 infections performed by fitting asymptotic distributions to actual data. By taking as a case-study the epidemic evolution of total COVID-19 infections in Chinese provinces and Italian regions, we find that predictions are characterized by large uncertainties at the early stages of the epidemic growth. Those uncertainties significantly reduce after the epidemics peak is reached. Differences in the uncertainty of the forecasts at a regional level can be used to highlight the delay in the spread of the virus. Our results warn that long term extrapolation of epidemics counts must be handled with extreme care as they crucially depend not only on the quality of data, but also on the stage of the epidemics, due to the intrinsically non-linear nature of the underlying dynamics. These results suggest that real-time epidemiological projections should include wide uncertainty ranges and urge for the needs of compiling high-quality datasets of infections counts, including asymptomatic patients.

Keywords: COVID-19, Logistic model, Epidemic model, National vs. Regional diffusion

1. Introduction

The COVID-19, a disease caused by the SARS-CoV-2 virus, was firstly reported in the Hubei province on 31 December 2019 when the WHO China Country Office was informed of cases of pneumonia unknown etiology detected in Wuhan City [1, 2, 3]. On 7 January 2020 the Chinese authorities identified this virus as a zoonotic virus belonging to the family of coronavirus [4, 5, 6]. Its diffusion rapidly spread over all Chinese provinces and nearest countries (Thailand, Japan, Korea) [7]. On 23 January, although still unknown the initial source of the epidemic, the evidence that 2019-nCoV spreads from human-to-human and also across generations of cases quickly increases [8, 9]. On 30 January, the World Health Organization (WHO) declared the outbreak to be a public health emergency of international concern [10], believing that it is still possible to interrupt the virus spread by putting in place strong measures for early detecting, isolating, and treating cases, for tracing back all contacts, and for promoting social distancing measures [10, 11, 12]. The main driver of transmission is still an open question [13, 14], and preliminary estimates of the median incubation period are 5-6 days (ranging between 2 and 14 days) [15]. On 21 February a cluster of cases was detected in Italy (Lombardia), then on 23 February 11 municipalities in northern Italy were identified as the two main Italian clusters and placed under quarantine [16], on 9 March the quarantine has been expanded to all of Italy [17], on 11 March all commercial activity except for supermarkets and pharmacies were prohibited [18], and on 22 March all non-essential businesses and industries were closed [19] and additional restrictions to movement of people were introduced [20, 21].

Meanwhile, the quarantined Chinese regions observed a fast decrease in the number of cases in Hubei and a moderate decrease in other affected regions, at the same time the virus internationally spread, and on 11 March the WHO declared COVID-19 a pandemic [22, 23]. To date, there are more than 1 million confirmed cases over the globe, more than 60000 deaths, and the most affected areas are the European region and the United States. While three months were needed to reach the first 100000 confirmed cases, only 23 days were sufficient to multiply by eight the counts, a typical signature of the exponential spreading of viruses. The reason for such high infectivity are currently being explored in clinical studies and numerical simulations [24]. Due to the fast spread of the virus and the severity of symptoms, restrictive confinement measures have been imposed in many countries. They were based

on asymptotic extrapolation of infection counts obtained on the basis of compartmental epidemic models as the Susceptible-Exposed-Infected-Recovered (SEIR) model and their variants [25] or on agent-based models [26]. Unfortunately, predictions made using these models are extremely sensitive to the underlying parameters and the quality of their extrapolation is deeply affected from both the lack of high-quality datasets as well as from the intrinsic sensitivity of the dynamics to initial conditions in the growing phase [27]. Moreover, in order to provide reliable estimates of asymptotic infection counts, a knowledge of asymptomatic populations is needed. These data are currently almost unavailable and affected by great uncertainties.

Another possibility is to extrapolate the number of infections by means of fitting asymptotic distributions to actual data. Using these phenomenological statistical approach, we compare the behavior of epidemic evolution across China and Italy. The assumption beyond those fits is that typical curves of total infections in SEIR models display a sigmoid shape [28]. Sigmoid functions such as the logistic or Gompertz can therefore be used to fit actual data. When data are collected with the same protocols, e.g., in China and Italy, where tests are performed only to symptomatic patients, the statistical fitting can therefore provide an extrapolation of how many symptomatic cases should be recorded, although it will not inform about the real percentage of infected population [29]. We found that predictions are characterized by large uncertainties at the early stages of the epidemic growth, significantly reducing when a mature stage or a peak of infections are reached. This is observed both in China and in Italy, although some differences are observed across the Italian territory, possibly related with the time delayed diffusion of epidemic into the different Italian regions. Finally, we also estimate infection increments for each Italian region, with being the uncertainty significantly reduced for Northern and Central regions, while a larger one is found for Southern regions. These results can be helpful for each epidemic diffusion, thus highlighting that the confinement measures are fundamental and more effective in the early stages of the epidemic evolution (the first 7 days), thus producing a different spread across provinces/regions as these measures are considered. **The main novelty introduced in this work is to investigate how uncertainty changes during the different stages of the epidemics. This is a crucial aspect that needs to be carefully considered when long-term extrapolations of the infection counts are carried out since they significantly depend not only on the quality of data, but also on the stage of the epidemics, due**

to the intrinsically non-linear nature of the underlying dynamics. This has also profound consequences on modeling epidemic growth by means of dynamical models as those based on compartments or agent dynamics which need to be initialized with quality data, faithfully representing the infected populations including asymptomatic patients [27]. Our approach, based on a sort of Bayesian framework to reduce uncertainty as more data and/or information become available, is particularly helpful for unknown viruses and outbreaks, and allows to suggest few practical guidelines to control the local diffusion of epidemics and to restrict the analysis on specific regions, aiming at preserving the public health and at enforcing/relaxing confinement measures.

2. Data

Data for the Chinese provinces are obtained from the data repository for the 2019 Novel Coronavirus Visual Dashboard operated by the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE), freely available at `https://github.com/CSSEGISandData/COVID-19`. Fig. 1 reports the total number of confirmed infections (left panel), thus including actual positive people to COVID-19, recovered and deaths for China and three Chinese provinces of Bejing, Hubei, and Yunnan, and the daily infections (right panel), during the period between 22 January and 30 March.

Data for the Italian regions are instead derived from the repository freely available at `https://github.com/pcm-dpc/COVID-19` where data are collected from the Italian Protezione Civile from 24 February 2020. Data used here were last downloaded on 02 April, thus covering the period 24 February - 02 April, as shown in Fig. 2. It is evident that although the increments of infections started about 1 month after the Chinese epidemic Italy has fast reached and exceeded the Chinese peak values of $\sim$80000 infections. Moreover, it is also apparent that epidemic diffusion in China reached its peak within $\sim$20 days from the first restriction operated to the Hubei region on 23 January. Conversely, the Italian restrictions seem to become more efficient only when the Italian government adopted a lock-down confinement on 9 March [17].

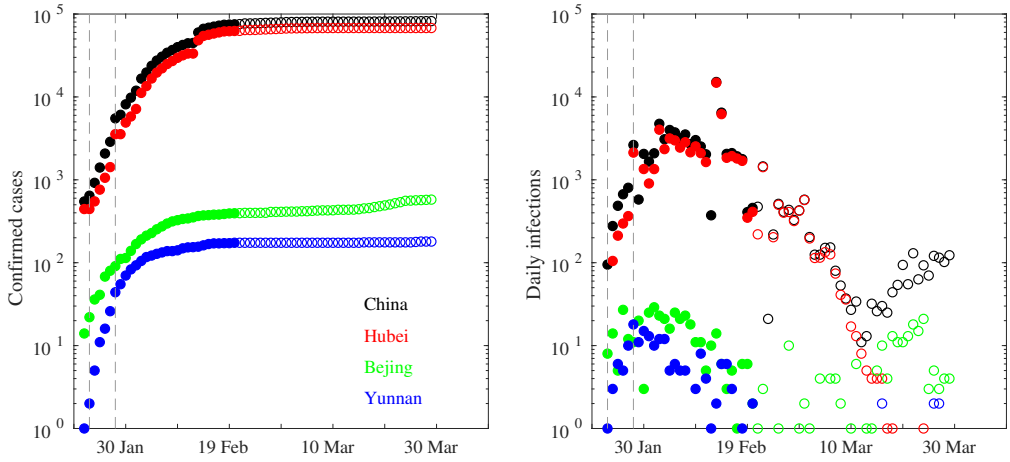

Figure 1: The total number of confirmed infections (left panel) and the daily infections (right panel) for China and three Chinese provinces of Bejing, Hubei, and Yunnan. Filled circles refer to the first 30 days of the epidemic diffusion. The vertical dashed lines mark the times when Chinese government applied lock-down restrictions on 23 January and 28 January, respectively.

3. Methods

A data-driven way to extrapolate future phases, **in terms of both key parameters and epidemic impact,** of an epidemic growth [30, 31, 32] is to use **phenomenological statistical models [33]. Indeed, since the total number of infections** $C(t)$ **is a sigmoid function different kinds of models can be used to fit its time evolution [34]. Within the large variety of possible sigmoid functions the generalized logistic distribution and the generalized Gompertz one have proven to be successful in fitting epidemic growths [35, 36]. Their suitability is mostly related to the reduced number of free parameters (only three) with respect to other choices depending on a larger set of model parameters (>3) which allows to reduce the overfitting effect due to a statistical model containing more parameters than can be justified by the data [34]. However, our main aim of investigating how uncertainty evolves with the epidemic growth stage is independent on the choice of the fitting distribution provided that they are dependent on the same number of free parameters.**

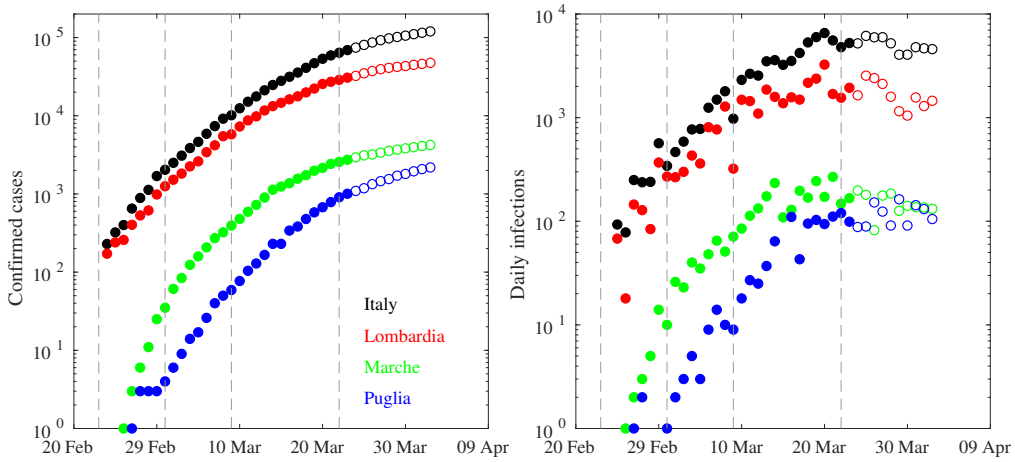

Figure 2: The total number of confirmed infections (left panel) and the daily infections (right panel) for Italy and three Italian regions of Lombardia, Marche, and Puglia. Filled circles refer to the first 30 days of the epidemic diffusion. The vertical dashed lines mark the times when the Italian government applied lock-down restrictions on 23 February, 01 March, 09 March, and 22 March, respectively.

Thus, we selected to use the generalized logistic distribution, also considering that its parameters can be linked (in a non-explicit way) to the solution of compartmental models as the Susceptible-Exposed-Infected-Recovered (SEIR) model and their variants [25] or on agent-based models [26]. The generalized logistic distribution for fitting the total cumulative number of infections reads [35, 33, 36]

$$C(t) = \frac{\alpha}{1 + \beta\, e^{-\gamma t}} \tag{1}$$

being α, β, and γ the parameters of the model. They can be fitted, e.g., using Nonlinear least-squares solver, with the Levenberg-Marquardt algorithm and the bisquare weight methods to minimize a weighted sum of squares. Here we use a MATLAB function to perform the fits. As recently pointed out in [27] in the early stages of the epidemics, the smoothness of COVID-19 cumulative infections data could lead to very uncertain predictions although with very good R^2. To avoid this, here we focus only on Chinese and Italian data, that, to date, represent a mature stage of the epidemics. This implies, as we will show, that the significance of the logistic fit can be assigned with greater

confidence [27]. We remark however, that when confinement measures are applied, the basic reproduction number R_0, which regulates the growth of infections, is reduced [37]. We are therefore in presence not of a single logistic distribution, but of a mixture of distributions with control parameters changing in time as different phases of epidemic diffusion are reached. Confinement measures can reduce R_0 from the exponential-like behavior of an uncontrolled growing phase, to a smoother logistic growth phase. Our goal here is to use the a-priori knowledge of the introduction of confinement measurements to investigate the perfomrance of statistical prediction of infection counts for different epidemic phases. Thus, we perform logistic fits as in Eq. (1) in the following time intervals:

- the first 30 days of epidemic growth, as reported in Figs. 3-4 by black lines, thus to consider how restrictions measure globally affect the diffusion;
- the first 7 days, roughly corresponding to the time interval during which first restriction measures are adopted both in China and Italy, although not still completely efficient (red lines in Figs. 3-4);
- the first 14 days, corresponding to the time interval in which the initial confinement measures should lead the first effects (blue lines in Figs. 3-4);
- the time interval between the 8^{th} and the 14^{th} day to investigate how the epidemic would be grown if starting from initial restrictions (green lines in Figs. 3-4);
- the time interval between the 15^{th} and the 30^{th} day to investigate the efficiency of restriction measures (magenta lines in Figs. 3-4).

In this way we can investigate both the efficiency of restriction measures in containing epidemic growth as well as the stability of prediction models based on logistic distribution fitting procedures. Moreover, to assess the significance of fits we assume that the last point of the fitting range could be affected by a $\pm 30\%$ error. This allows us to provide a simple way to estimate confidence intervals for our fits [27]. Finally, the Kolmogorov-Smirnov (K-S) test [38, 39, 40] is also used to obtain a test decision for the null hypothesis that the observed data are from the same logistic distribution as derived from the logistic fits under the different time intervals. This allows to test

the efficiency in delivering reliable forecasts at different stages of the epidemic growth. The test is based on evaluating the maximum distance between the empirical distribution functions coming from two different samples $x_{1,n}$ and $x_{2,m}$, being n and m the length of samples. By defining the Kolmogorov-Smirnov statistic as

$$D_{n,m} = \sup_x |F_{1,n}(x) - F_{2,m}(x)|, \tag{2}$$

where $F_{1,n}(x)$ and $F_{2,m}(x)$ are the empirical distribution functions of the two samples, respectively, the null hypothesis is rejected at the confidence level α if

$$D_{n,m} > c(\alpha)\sqrt{\frac{n+m}{n \cdot m}}. \tag{3}$$

When $m = n$ a general relation can be found for $D_n(\alpha)$ as

$$D_n(\alpha) > \frac{1}{\sqrt{n}}\sqrt{-\log\left(\frac{\alpha}{2}\right)}. \tag{4}$$

The value of $c(\alpha)$ for the most common levels of α are reported in Table 1.

α	0.20	0.15	0.10	0.05	0.01
$c(\alpha)$	1.073	1.138	1.224	1.358	1.628

Table 1: The value of $c(\alpha)$ for the most common levels of α.

The closer the observed statistics $D_{n,obs}$ is to 0 the more likely it is that the two samples were drawn from the same distribution with being $D_{n,obs} < D_n(\alpha)$. The use of the K-S test has two main advantages: i) the distribution of the K-S test statistic itself does not depend on the underlying cumulative distribution function being tested, and ii) it is an exact test [41, 42, 43, 44]. Moreover, it is specifically designed for testing if data come from a normal, a log-normal, a Weibull, an exponential, or a logistic distribution [42, 45]. Thus, it is particularly suitable for our investigations, being also a non-parametric and robust technique since it is not based on strong distributional assumptions [42, 45, 46, 44].

4. Epidemic diffusion through Chinese provinces

Fig. 3 shows logistic fits for different phases of epidemic across Chinese provinces, together with upper and lower confidence bounds, obtained as out-

lined in the previous section. Early stage of epidemic propagation is characterized by a larger confidence interval (red lines in Fig. 3), thus highlighting the difficulty in making early reliable predictions of epidemic growth, with an exponential-like behavior. The confidence interval becomes narrower as the growth rate reduces, as for the case of the provinces of Bejing and Yunnan being less affected from COVID-19 infections with respect to the Hubei, the latter mostly contributing to the overall epidemic growth in China. The logistic fit becomes more stable, being characterized by a narrower estimates of confidence intervals, when the first two weeks are considered (blue lines in Fig. 3), possibly related to the initial efficiency of restriction measures. This could be also due to both the limited number of points of the fitting range as well as to the particular phase of the epidemic growth. However, by comparing the confidence intervals of logistic fits performed using the first week (22/01 - 29/01, red lines in Fig. 3) and the second week (30/01 - 05/02, green lines in Fig. 3) it is possible to note that the stability increases for this second interval for all Chinese provinces, thus suggesting that estimates are significantly dependent on the particular epidemic phase considered. Indeed, the stability significantly increases when the logistic fit is performed on time intervals that do not include the first week of the epidemic growth (green and magenta lines in Fig. 3), suggesting that credible predictions could be assigned with a large confidence by means of a logistic fit if the beginning of the outbreak is not considered. However, the narrowest estimates of significance levels is obtained when the first 30 days are considered, thus also including the beginning of the outbreak, possibly suggesting that fits become more and more stable if data are collected at a mature stage of the epidemic growth. This is clearly visible for all Chinese provinces, apart for the slight discrepancy observed for the Bejing province where some returned cases from outside China were observed from 20 March. Finally, we assess the statistical discrepancy of the logistic fits from the observed data by performing the Kolmogorov-Smirnov (K-S) test those results for the 95% confidence level are reported in Table 2.

It can be noted that the statistical results obtained through the K-S test suggest that the fits performed by considering the time intervals from 22 January to 21 February as well as from 05 February to 21 February are statistically significant for reproducing the behavior of the observed number of infections at the 95% significance level. This seems to support the that reliable predictions can be assessed only when a mature stage of the epidemic growth is approached/reached, while low-significant predictions can

	$D_{n,obs}$			
Time interval	China	Hubei	Bejing	Yunnan
22/01 - 29/01	0.750	0.750	0.750	0.625
22/01 - 05/02	0.500	0.475	0.550	0.450
30/01 - 05/02	0.575	0.575	0.550	0.525
05/02 - 21/02	**0.225**	**0.150**	**0.125**	**0.125**
22/01 - 21/02	**0.100**	**0.100**	**0.100**	**0.100**

Table 2: Results of the Kolmogorov-Smirnov test for the 95% confidence level for the Chinese provinces. The decision to reject the null hypothesis is based on comparing the observed statistics $D_{n,obs}$ with the theoretical value $D_{n,th} = 0.2329$ obtained for the significance level $\alpha = 0.05$ as in Eq. 4. If $D_{n,obs} < D_{n,th}$ then the samples come from the same logistic distribution and corresponding values are reported in bold.

be released at the early stages of the epidemic diffusion.

5. Epidemic diffusion through Italian regions

Fig. 4 shows logistic fits for different phases of epidemic across Italian regions, together with the upper and lower confidence lines. As for Chinese provinces the early stage of epidemic diffusion is characterized by a larger confidence interval (red lines in Fig. 4), again suggesting that reliable predictions of epidemic growth are particularly difficult in its early stages. Indeed, an exponential-like behavior is found for both the Italian territory and Lombardia, the latter being the first Italian region characterized by COVID-19 infections. As for China, confidence intervals become narrower as the growth rate reduces (see for example Marche or Puglia with respect to Lombardia), with the logistic fits also becoming more stable when the initial stages of the outbreak are removed (green and magenta lines in Fig. 4). Unlike for Chinese regions, Italian regions present a wide range of different epidemic behaviors, that we investigate separately in the following.

5.1. Epidemics growth in Lombardia

As discussed above the initial phase is characterized by larger uncertainties and by an exponential-like behavior (red lines in Fig. 4), thus suggesting a clear difficulty in making predictions of the growth in its early stage. When the first two weeks (e.g., 24/02 - 08/03) are considered (blue lines in Fig. 4) a larger uncertainty is found, especially for the upper-bound confidence level. This underline the difficulty in making reliable estimates of its evolution.

Similarly, the logistic fits performed between 01 March and 08 March (green lines in Fig. 4) suggest that the first two weeks were particularly critical in Lombardia, while logistic fits become more stable when removing the beginning of the outbreak, leading to more reliable estimates of the epidemic growth (magenta lines in Fig. 4). Finally, significance levels become narrower when the first 30 days are considered (e.g., 24/02 - 23/03), thus also including the beginning of the outbreak, possibly again suggesting that including data from the mature stage of the epidemic growth could allow to obtain more stable fits. We remark that, no matter the approach followed, logistic fits struggle to predict the number of infections of the successive days. This failure of statistical real-time forecasts of the epidemics could be related to all those factors that can change the instantaneous value of R_0, e.g., extended violations of the restriction measures, changes in testing protocols or delay in data reporting, changes in the virus characteristics. It is worthwhile to note that the above features are found for all Northern regions firstly affected from COVID-19 diffusion (not shown here).

5.2. Epidemics growth in Marche

The epidemic growth throughout Marche, as well as throughout other Central regions (not shown), is different from Northern regions. Indeed, the first 7 days (e.g., 24/02 - 01/03, red lines in Fig. 4) were not characterized by an exponential increase of infections, as the diffusion of the virus was pretty slow: logistic fits are therefore meaningless in this context. The exponential phase started in the second week, as we can see by fitting the first two week of the infection counts (e.g., 24/02 - 08/03, blue lines in Fig. 4) or just the second week (e.g., from 01 March to 08 March, green lines in Fig. 4). During this week, the number of infections significantly increases (272 confirmed cases) enabling better fits of data to logistic distributions. This suggests a time delayed propagation between Northern and Central regions. Indeed, the logistic fits become more stable, with narrower estimates of confidence intervals, when the time interval from 08 March to 23 March (magenta lines in Fig. 4) or the first 30 days (e.g., 24/02 - 23/03, black lines in Fig. 4) are taken into account, suggesting that credible predictions could be assigned with a large confidence when a mature stage of the epidemic growth is approached. However, as for Norther regions the logistic fits struggle to predict the number of infections of the successive days (i.e., after the first 30 days).

5.3. Epidemics growth in Puglia

A completely different scenario is found for Puglia and Southern regions (not shown). Logistic fits cannot be performed during during the first two weeks (e.g., from 24 February to 08 March), as the infection counts was not yet exponential. By considering the time interval between 08 and 23 March (magenta lines in Fig. 4) and the first 30 days (e.g., 24/02 - 23/03, black lines in Fig. 4) an increase in the confidence of logistic fits is found, although they struggle to predict the number of infections of the successive days (i.e., after the first 30 days). This is possibly due to the time delayed propagation of epidemic throughout Southern regions for which a mature stage is, to date, not yet reached. To support this hypothesis and to assess the statistical discrepancy of the logistic fits from the observed data we perform the Kolmogorov-Smirnov (K-S) test those results for the 95% confidence level are reported in Table 3.

	$D_{n,obs}$			
Time interval	Italy	Lombardia	Marche	Puglia
24/02 - 01/03	0.825	0.800	0.800	0.800
24/02 - 08/03	0.575	0.550	0.650	0.800
01/03 - 08/03	0.550	0.425	0.600	0.800
08/03 - 23/03	0.325	0.325	0.400	0.400
24/02 - 23/03	0.350	0.325	0.400	0.400

Table 3: Results of the Kolmogorov-Smirnov test for the 95% confidence level for the Italian regions. The decision to reject the null hypothesis is based on comparing the observed statistics $D_{n,obs}$ with the theoretical value $D_{n,th} = 0.3037$ obtained for the significance level $\alpha = 0.05$ as in Eq. 4. If $D_{n,obs} < D_{n,th}$ then the samples come from the same logistic distribution and corresponding values are reported in bold.

It is interesting to note that, although lower values of $D_{n,obs}$ are observed when a more mature stage of the epidemic growth is considered in the fitting range, as for example for time intervals from 24 February to 23 March as well as from 08 to 23 March, the observed values $D_{n,obs}$ are all above the statistical threshold of $D_{n,th} = 0.3037$. This suggests that a mature stage is, to the date of 23 March, not yet reached, although Northern and Central regions are characterized by lower values than the Southern ones, thus possibly related to the time delayed propagation of epidemic throughout Southern regions.

6. Estimation of infections for Italy and their peak time

As discussed in Section 5 all performed logistic fits struggle to predict the number of infections of the successive days (i.e., after the first 30 days), thus we performed and compare logistic fits in three time intervals: (i) the first 30 days (e.g., from 24 February to 23 March), (ii) the first 37 days (e.g., from 24 February to 30 March), and (iii) the overall period from 24 February to 02 April. The results of the Kolmogorov-Smirnov test for the 95% confidence level are reported in Table 4, while the behavior of logistic fits are shown in Fig. 5.

	$D_{n,obs}$			
Time interval	Italy	Lombardia	Marche	Puglia
24/02 - 23/03	0.350	0.325	0.400	0.400
24/02 - 30/03	**0.150**	**0.150**	**0.250**	**0.275**
24/02 - 02/04	**0.100**	**0.100**	**0.175**	**0.200**

Table 4: Results of the Kolmogorov-Smirnov test for the 95% confidence level for the Italian regions. The decision to reject the null hypothesis is based on comparing the observed statistics $D_{n,obs}$ with the theoretical value $D_{n,th} = 0.3037$ obtained for the significance level $\alpha = 0.05$. If $D_{n,obs} < D_{n,th}$ then the samples come from the same logistic distribution.

It is interesting to note that all regions and Italy are characterized by lower values of $D_{n,obs}$, below the theoretical value $D_{n,th} = 0.3037$, when including the next 7 days (e.g., by considering the period between 24 February and 30 March) to the logistic fits and when considering the whole time range (e.g., 24/02 - 02/04). Lombardia presents lower values of the K-S statistics $D_{n,obs}$ than those for Marche and Puglia, together with a narrower confidence interval when including the successive days, not observed for both Marche and Puglia. Particularly for Puglia the confidence interval remains practically unchanged, thus suggesting that logistic fits are not still stable, possibly due to the fact that Southern regions have not yet reached a mature stage of the epidemic growth. This difference in terms of stability of logistic fits as well as on confidence of reliable estimates can be clearly seen by looking at the behavior of estimated daily increments. Days of peak significantly depends on the fitting range for Puglia, while the estimation of this quantity is more stable for Lombardia and Marche, as shown in Fig. 6. Indeed a wider discrepancy is found between daily increments

and estimates for logistic fits performed during the three intervals, obviously affecting both the peak time estimation and its value. By comparing our estimates and data collected from the daily report of the Italian Protezione Civile (`https://github.com/pcm-dpc/COVID-19`) we found that the discrepancy significantly increases when moving from Northern to Southern regions, where it can also reach an error which is comparable with the predicted value. This could be the reflection of at least two different factors: i) the epidemic growth is in a more mature phase in the Northern and Central regions with respect to the Southern ones, where it began with a time delay ranging from 3 to 14 days, and ii) the higher ratio between the observed cases and the number of tests carried out for Southern regions with respect to the rest of Italy (see `https://github.com/pcm-dpc/COVID-19`). These two factors could affect the performance of the logistic fits for the Southern regions of Italy, being characterized by wider uncertainties with respect to the rest of Italy. Thus, our results suggest that estimates of the ending of epidemic growth are affected by the statistical uncertainties, by the delayed propagation of infections through the different regions, and by the effective respect of the guidelines in terms of confinement measures.

7. Conclusion

In this paper we investigated the behavior of predictions of COVID-19 infections on the particular phase of its growth and propagation in a specific country, as well as, on the effectiveness of social distancing and confinement measures. By analyzing the epidemic evolution in China and Italy we find that predictions are characterized by large uncertainties at the early stages of the epidemic growth, significantly reducing when the epidemics peak is past, independently on how this is reached. While infection counts for different Chinese provinces show a synchronised behavior, counts for Italian regions point to different epidemic phases. While the epidemic peak has been likely reached in the Northern and Central regions, COVID-19 infections are still in a growing phase for Southern regions, with a delay ranging from 3 to 14 days. By assessing the performance of logistic fits we assess that a wider uncertainty is found during the first week of epidemic propagation. Uncertainty is reduced when data from the very beginning of the breakout are removed from the datasets. Moreover, the estimated infection increments are extremely sensitive to the epidemic growth stage and to the last points considered to perform statistical extrapolations. Higher significance levels

are reached for the more mature stages of the epidemic growth.

The most interesting pattern in the time-evolution of the distribution is the observed change from an exponential-like behavior observed at the beginning of the epidemic growth to a sigmoid-like one when first restriction measures are introduced, particularly evident for the Italian case study. Indeed, by evaluating the expected final number of total infections as predicted from logistic fits during the different stages we highlight that reliable estimates cannot be released until more mature stages of the epidemic growth are reached. We show that by only means of the first 7 days, corresponding to the time interval during which first restriction measures are adopted both in China and Italy, an overestimation of the final number of infections of ∼65% for China and ∼2000% for Italy is observed. Conversely, by considering the first 14 days, corresponding to the time interval in which the initial confinement measures should lead the first effects, an underestimation of ∼-48% for China and ∼-76% for Italy is obtained. A lower underestimation (∼-32% and ∼-69% for China and Italy, respectively) is found when considering the time interval between the 8^{th} and the 14^{th} day, e.g., by investigating how the epidemic would be grown if starting from initial restrictions only, while a better agreement is found when considering the time interval between the 15^{th} and the 30^{th} day, corresponding to investigate the efficiency of restriction measures, with reduced underestimation of the final number of infections of ∼-17% for China and ∼-12% for Italy. Finally, by monitoring the stability of logistic fits as well as their suitability on predicting the number of infections of the successive days (i.e., after the first 30 days) we highlight how the uncertainty evolution can be used to track how the epidemics diffused at a regional level, allowing an estimation of the delay in the spread of the virus. Indeed, we found that the uncertainty significantly increases when moving from Northern to Southern regions, where the error is almost comparable with the predicted value, suggesting that, to date, the epidemic peak has not been likely reached for Southern regions, being delayed with respect to Northern and Central ones.

Our results aim at providing some guidelines for real-time epidemics forecasts which should be applicable to other viruses and outbreaks. Real-time forecasts of the epidemics are, to date, a societal need more than a scientific

field. They are crucial to plan the duration of confinement measures and to define the needs for health-care facilities. The aim of this letter was to show that those extrapolations crucially depend not only on the quality of data, but also on the stage of the epidemics, due to the intrinsically non-linear nature of the underlying dynamics. This prevents from performing successful long-term extrapolations of the infection counts with statistical models. **As a guideline it is surely helpful to perform logistic fits every day and to evaluate the reliability on predicting the next day, and then perform a new logistic fit to investigate how the uncertainty grown/reduced. Moreover, reliable estimates are surely affected by possible source of errors in counting infections, thus we suggest to assess the significance of fits to the last data point of the fitting range by assuming it could be affected by a $\pm$30% error. This allows us to provide a simple way to estimate confidence intervals [27]. Furthermore, we also suggest not only to exclude the last data point and check fits stability but also to consider to exclude initial point(s) to evaluate how epidemic would be grown if starting from initial restriction measures or how delayed propagation could be present at a regional level.**

Our approach, based on a sort of Bayesian framework to update the probability for a reduced uncertainty as more evidence or information become available (this especially true for unknown viruses and outbreaks), suggests that the statistical modeling of epidemic growth should be focused on specific stages of its evolution on time as well as on its spread at a more local level (e.g., regional level). This can help in controlling local diffusion of epidemics and to restrict the analysis on specific regions depending on its uncertainty values. The above guidelines can be also suitable for dynamical models such as those based on compartments or agent dynamics which need to be initialized with quality data, faithfully representing the infected populations including asymptomatic patients [27]. It is therefore crucial to pursue national health systems to provide the most transparent and extended datasets as possible and obtain high quality datasets to initialize those models. We remind that only dynamical models can provide a coherent representation and evolution of the epidemics, as they are effectively based on the conservation of the total number of individuals. **Characterizing and modeling the uncertainty can allow to preserve the public health and help to enforce/relax strict confinement measures.**

Acknowledgments

The authors thank the anonymous reviewer for fruitful and insightful comments.

References

[1] R. F. Service, Coronavirus epidemic snarls science worldwide, Science 367 (2020) 836–837. doi:10.1126/science.367.6480.836.

[2] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, et al., Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China, The Lancet 395 (2020) 497–506. doi:10.1016/S0140-6736(20)30183-5.

[3] W.-j. Guan, Z.-y. Ni, Y. Hu, W.-h. Liang, C.-q. Ou, J.-x. He, L. Liu, H. Shan, C.-l. Lei, D. S. Hui, B. Du, L.-j. Li, G. Zeng, K.-Y. Yuen, R.-c. Chen, C.-l. Tang, T. Wang, P.-y. Chen, J. Xiang, S.-y. Li, J.-l. Wang, Z.-j. Liang, Y.-x. Peng, L. Wei, Y. Liu, Y.-h. Hu, P. Peng, J.-m. Wang, J.-y. Liu, Z. Chen, G. Li, Z.-j. Zheng, S.-q. Qiu, J. Luo, C.-j. Ye, S.-y. Zhu, N.-s. Zhong, Clinical characteristics of coronavirus disease 2019 in china, New England Journal of Medicine 0 (0) null. doi:10.1056/NEJMoa2002032.

[4] W. H. Organization, Pneumonia of unknown cause – china, 2020. URL: https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/.

[5] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li, T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang, S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao, Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Lam, J. T. Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, Z. Feng, Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia, New England Journal of Medicine 382 (2020) 1199–1207. doi:10.1056/NEJMoa2001316.

[6] M. Lipsitch, D. L. Swerdlow, L. Finelli, Defining the epidemiology of covid-19 — studies needed, New England Journal of Medicine 382 (2020) 1194–1196. doi:10.1056/NEJMp2002125.

[7] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, W. Tan, A novel coronavirus from patients with pneumonia in china, 2019, New England Journal of Medicine 382 (2020) 727–733. doi:10.1056/NEJMoa2001017.

[8] D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang, H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, X. Wang, Z. Peng, Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China, JAMA 323 (2020) 1061–1069. URL: https://doi.org/10.1001/jama.2020.1585. doi:10.1001/jama.2020.1585. arXiv:https://jamanetwork.com/journals/jama/articlepdf/2761044/jama_wang_2020_c

[9] L. T. Phan, T. V. Nguyen, Q. C. Luong, T. V. Nguyen, H. T. Nguyen, H. Q. Le, T. T. Nguyen, T. M. Cao, Q. D. Pham, Importation and human-to-human transmission of a novel coronavirus in vietnam, New England Journal of Medicine 382 (2020) 872–874. doi:10.1056/NEJMc2001272, pMID: 31991079.

[10] W. H. Organization, Coronavirus disease 2019 (covid-19) situation report-10, 2020. URL: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/.

[11] W. E. Parmet, M. S. Sinha, Covid-19 — the law and limits of quarantine, New England Journal of Medicine 0 (0) null. doi:10.1056/NEJMp2004211.

[12] R. L. Haffajee, M. M. Mello, Thinking globally, acting locally — the u.s. response to covid-19, New England Journal of Medicine 0 (0) null. doi:10.1056/NEJMp2006740.

[13] S. Flaxman, et al., Estimating the number of infections and the impact of non-pharmaceutical interventions on covid-19 in 11 european countries, 2020. URL: https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-Europe-estimates-and-NPI-impact-30-03-2020.pdf.

[14] C. Rothe, M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, C. Wallrauch, T. Zimmer, V. Thiel, C. Janke, W. Guggemos, M. Seilmaier, C. Drosten, P. Vollmar, K. Zwirglmaier, S. Zange, R. Wölfel, M. Hoelscher, Transmission of 2019-ncov infection from an asymptomatic contact in germany, New England Journal of Medicine 382 (2020) 970–971. doi:10.1056/NEJMc2001468.

[15] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, J. Lessler, The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application, Annals of Internal Medicine (2020). URL: https://doi.org/10.7326/M20-0504. doi:10.7326/M20-0504. arXiv:https://annals.org/acp/content_public/journal/aim/0/aime202005050-m20050

[16] P. M. of Italy, Disposizioni attuative del decreto-legge 23 febbraio 2020, n. 6, recante misure urgenti in materia di contenimento e gestione dell'emergenza epidemiologica da covid-19, 2020. URL: https://www.gazzettaufficiale.it/eli/id/2020/02/23/20A01228/sg.

[17] P. M. of Italy, Ulteriori disposizioni attuative del decreto-legge 23 febbraio 2020, n. 6, recante misure urgenti in materia di contenimento e gestione dell'emergenza epidemiologica da covid-19, applicabili sull'intero territorio nazionale, 2020. URL: https://www.gazzettaufficiale.it/eli/id/2020/03/09/20A01558/sg.

[18] P. M. of Italy, Ulteriori disposizioni attuative del decreto-legge 23 febbraio 2020, n. 6, recante misure urgenti in materia di contenimento e gestione dell'emergenza epidemiologica da covid-19, applicabili sull'intero territorio nazionale, 2020. URL: https://www.gazzettaufficiale.it/eli/id/2020/03/11/20A01605/sg.

[19] P. M. of Italy, Ulteriori disposizioni attuative del decreto-legge 23 febbraio 2020, n. 6, recante misure urgenti in materia di contenimento e gestione dell'emergenza epidemiologica da covid-19, applicabili sull'intero territorio nazionale, 2020. URL: https://www.gazzettaufficiale.it/eli/id/2020/03/22/20A01807/sg.

[20] L. Rosenbaum, Facing covid-19 in italy — ethics, logistics, and therapeutics on the epidemic's front line, New England Journal of Medicine 0 (0) null. doi:10.1056/NEJMp2005492.

[21] M. Chinazzi, J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. Pastore y Piontti, K. Mu, L. Rossi, K. Sun, C. Viboud, X. Xiong, H. Yu, M. E. Halloran, I. M. Longini, A. Vespignani, The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak, Science (2020). URL: https://science.sciencemag.org/content/early/2020/03/05/science.aba9757. doi:10.1126/science.aba9757. arXiv:https://science.sciencemag.org/content/early/2020/03/05/science.aba9757.

[22] A. S. Fauci, H. C. Lane, R. R. Redfield, Covid-19 — navigating the uncharted, New England Journal of Medicine 382 (2020) 1268–1269. doi:10.1056/NEJMe2002387.

[23] W. H. Organization, Coronavirus disease 2019 (covid-19) situation report-51, 2020. URL: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf?sfvrsn=1ba62e57_10.

[24] A. Spinello, A. Saltalamacchia, A. Magistrato, Is the Rigidity of SARS-CoV-2 Spike Receptor-Binding Motif the Hallmark for Its Enhanced Infectivity and Pathogenesis? (2020). URL: https://chemrxiv.org/articles/Is_the_Rigidity_of_SARS-CoV-2_Spike_Receptor-Binding_Motif_the_Hallmark_for_Its_Enhanced_Infectivity_and_Pathogenesis_/12091260. doi:10.26434/chemrxiv.12091260.v1.

[25] F. Brauer, Compartmental models in epidemiology, Lecture Notes in Mathematics, Springer Verlag, 2008, pp. 19–79.

[26] S. L. Chang, N. Harding, C. Zachreson, O. M. Cliff, M. Prokopenko, Modelling transmission and control of the covid-19 pandemic in australia, arXiv preprint arXiv:2003.10218 (2020).

[27] D. Faranda, I. P. Castillo, O. Hulme, A. Jezequel, J. S. W. Lamb, Y. Sato, E. L. Thompson, Asymptotic estimates of sars-cov-2 infection counts and their sensitivity to stochastic perturbation, Chaos under review (2020).

[28] T. C. Wilcosky, L. E. Chambless, A comparison of direct adjustment and regression adjustment of epidemiologic measures, Journal of chronic diseases 38 (1985) 849–856.

[29] M. Batista, Estimation of the final size of the covid-19 epidemic, Preprint.] medRxiv (2020).

[30] C.-H. Li, C.-C. Tsai, S.-Y. Yang, Analysis of the permanence of an SIR epidemic model with logistic process and distributed time delay, Communications in Nonlinear Science and Numerical Simulations 17 (2012) 3696–3707. doi:10.1016/j.cnsns.2012.01.018.

[31] J. Kumar, K. P. S. S. Hembram, Epidemiological study of novel coronavirus (COVID-19), arXiv e-prints (2020) arXiv:2003.11376. arXiv:2003.11376.

[32] J. Ma, Estimating epidemic exponential growth rate and basic reproduction number, Infectious Disease Modelling 5 (2020) 129 – 141. URL: http://www.sciencedirect.com/science/article/pii/S2468042719300491. doi:https://doi.org/10.1016/j.idm.2019.12.009.

[33] G. Chowell, Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts, Infectious Disease Modelling 2 (2017) 379 – 398. URL: http://www.sciencedirect.com/science/article/pii/S2468042717300234. doi:https://doi.org/10.1016/j.idm.2017.08.001.

[34] I. J. Wellock, G. C. Emmans, I. Kyriazakis, Describing and predicting potential growth in the pig, Animal Science 78 (2004) 379–388. doi:10.1017/S1357729800058781.

[35] B. Gompertz, On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies, Philosophical Transactions of the Royal Society of London Series I 115 (1825) 513–583.

[36] R. Bürger, G. Chowell, L. Y. Lara-Díiaz, Comparative analysis of phenomenological growth models applied to epidemic outbreaks, Mathematical biosciences and engineering : MBE 16 (2019) 4250—4273. URL: https://doi.org/10.3934/mbe.2019212. doi:10.3934/mbe.2019212.

[37] G. Viceconte, N. Petrosillo, Covid-19 r0: Magic number or conundrum?, Infectious Disease Reports 12 (2020). URL: https://

www.pagepress.org/journals/index.php/idr/article/view/8516. doi:10.4081/idr.2020.8516.

[38] A. N. Kolmogorov, Sulla determinazione empirica di una legge di distribuzione, Giornale dell'Istituto Italiano degli Attuari 4 (1933) 83–91.

[39] N. Smirnov, Table for estimating the goodness of fit of empirical distributions, Ann. Math. Statist. 19 (1948) 279–281. URL: https://doi.org/10.1214/aoms/1177730256. doi:10.1214/aoms/1177730256.

[40] M. A. Stephens, Edf statistics for goodness of fit and some comparisons, Journal of the American Statistical Association 69 (1974) 730–737. doi:10.1080/01621459.1974.10480196.

[41] T. W. Anderson, D. A. Darling, A test of goodness of fit, Journal of the American Statistical Association 49 (1954) 765–769. doi:10.1080/01621459.1954.10501232.

[42] G. Marsaglia, W. W. Tsang, J. Wang, Evaluating kolmogorov's distribution, Journal of Statistical Software, Articles 8 (2003) 1–4. URL: https://www.jstatsoft.org/v008/i18. doi:10.18637/jss.v008.i18.

[43] Kolmogorov–Smirnov Test, Springer New York, New York, NY, 2008, pp. 283–287. URL: https://doi.org/10.1007/978-0-387-32833-1_214. doi:10.1007/978-0-387-32833-1_214.

[44] M. Marozzi, Nonparametric simultaneous tests for location and scale testing: A comparison of several methods, Communications in Statistics - Simulation and Computation 42 (2013) 1298–1317. doi:10.1080/03610918.2012.665546.

[45] F. J. Massey, Distribution table for the deviation between two sample cumulatives, Ann. Math. Statist. 23 (1952) 435–441. URL: https://doi.org/10.1214/aoms/1177729388. doi:10.1214/aoms/1177729388.

[46] W. Ghidey, E. Lesaffre, G. Verbeke, A comparison of methods for estimating the random effects distribution of a linear mixed model, Statistical Methods in Medical Research 19 (2010) 575–600.

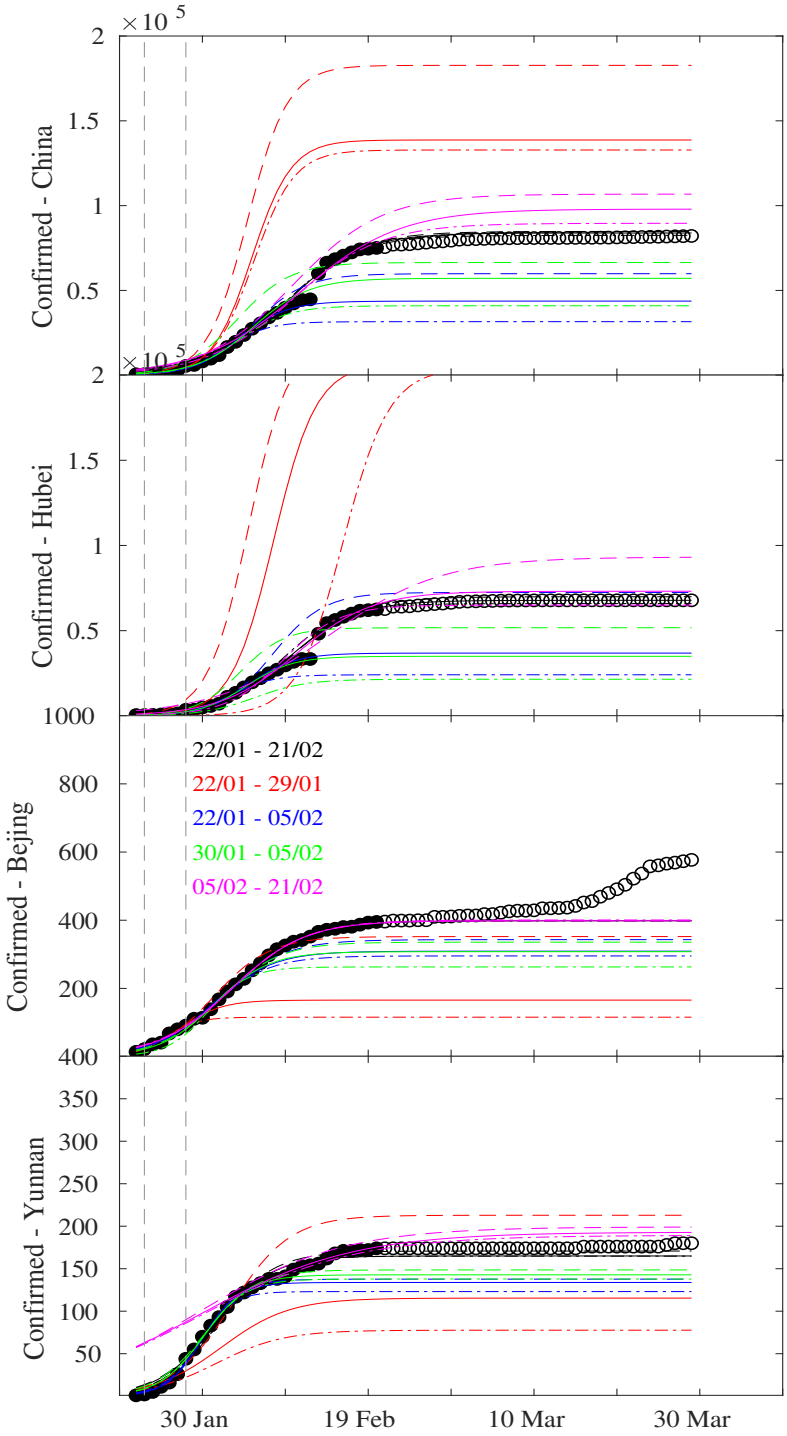

Figure 3: Logistic fits during the different time intervals of epidemic across Chinese provinces, together with the confidence lines. From top to bottom: China and three provinces (Bejing, Hubei, Yunnan). The vertical dashed lines mark the times when Chinese government applied lock-down restrictions on 23 January and 28 January, respectively.

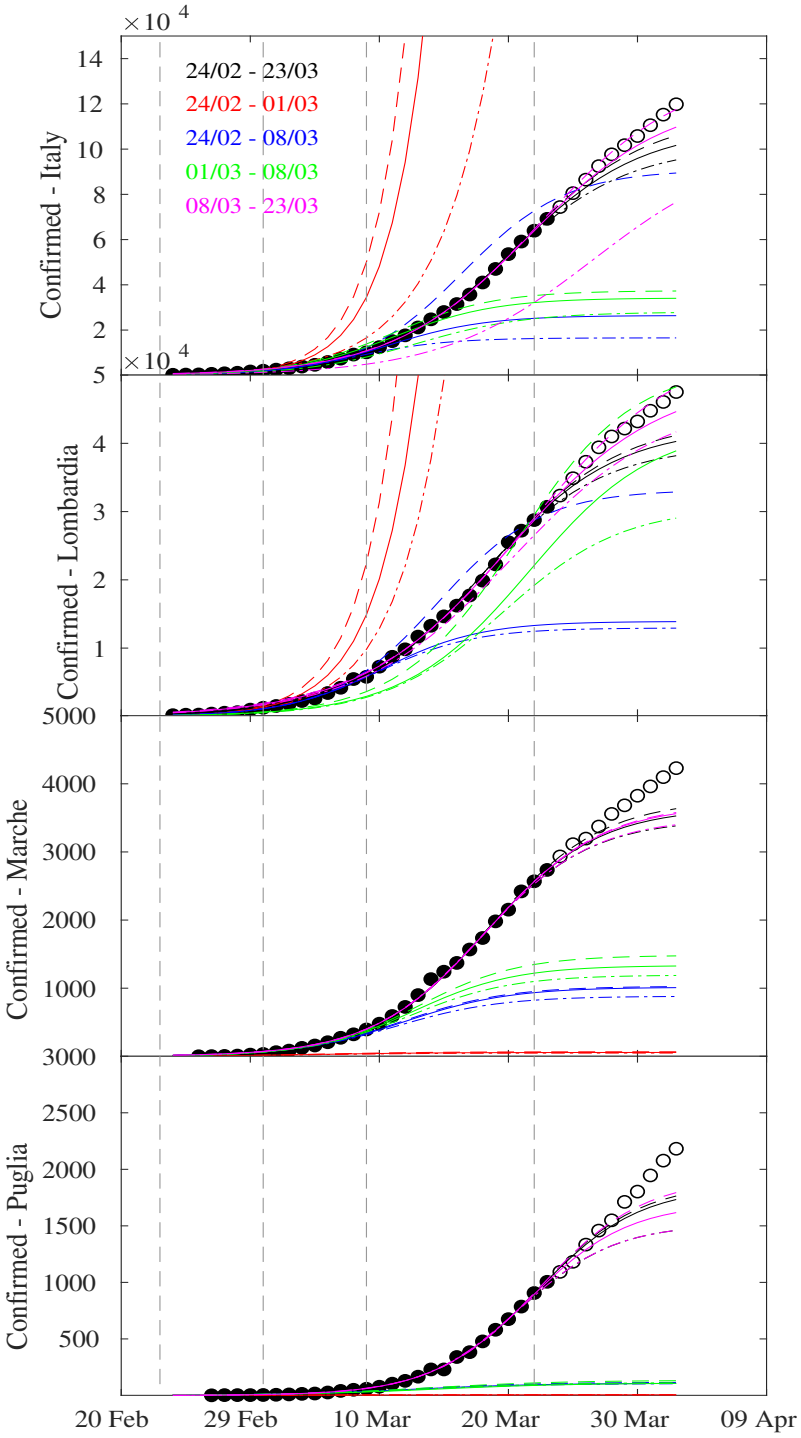

Figure 4: Logistic fits during the different time intervals of epidemic across Italian regions, together with the confidence lines. From top to bottom: Italy and three regions (Lombardia, Marche, Puglia). The vertical dashed lines mark the times when Italian government applied lock-down restrictions on 23 February, 01 March, 09 March, and 22 March, respectively.

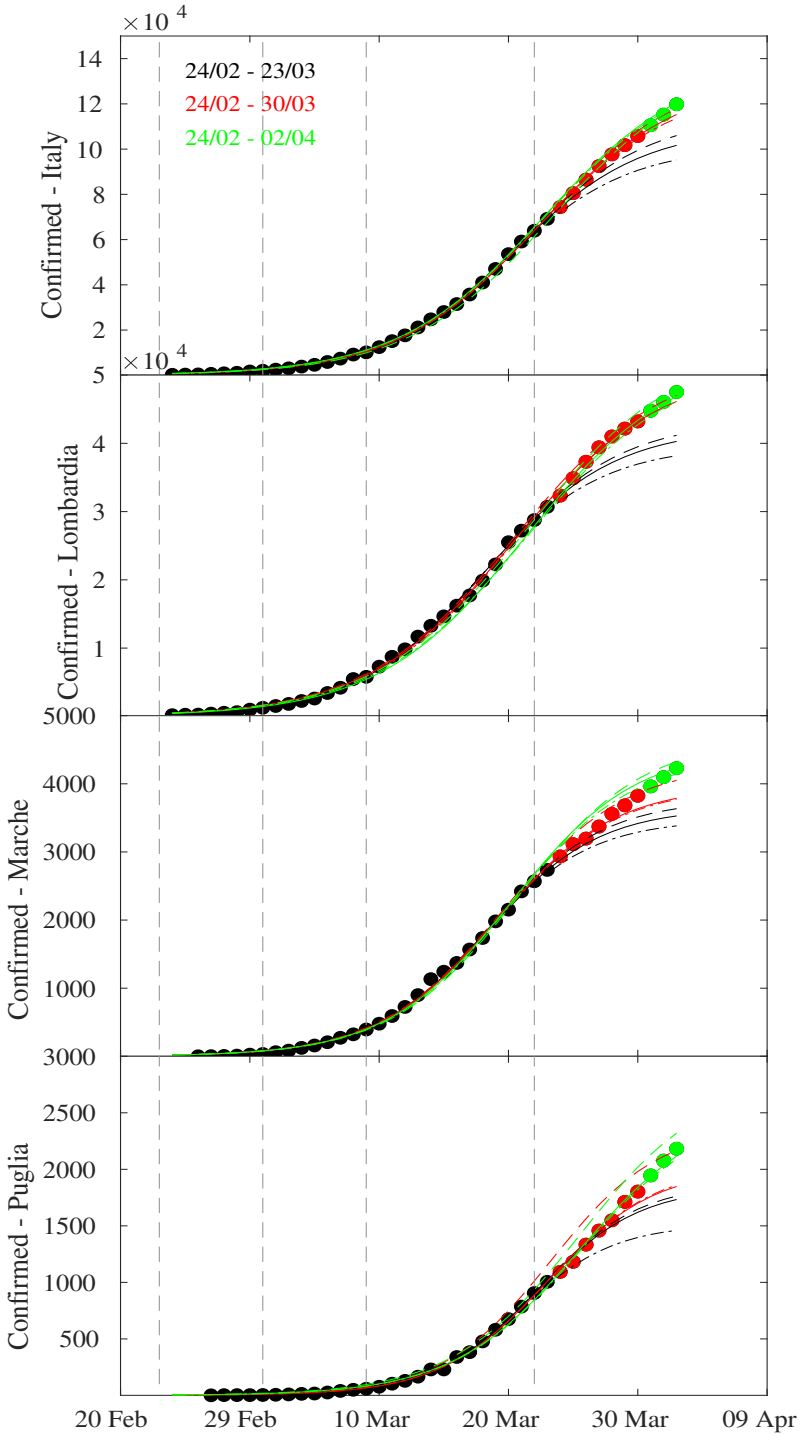

Figure 5: Logistic fits during the different time intervals of epidemic across Italian regions, together with the confidence lines. From top to bottom: Italy and three regions (Lombardia, Marche, Puglia). The vertical dashed lines mark the times when Italian government applied lock-down restrictions on 23 February, 01 March, 09 March, and 22 March, respectively.

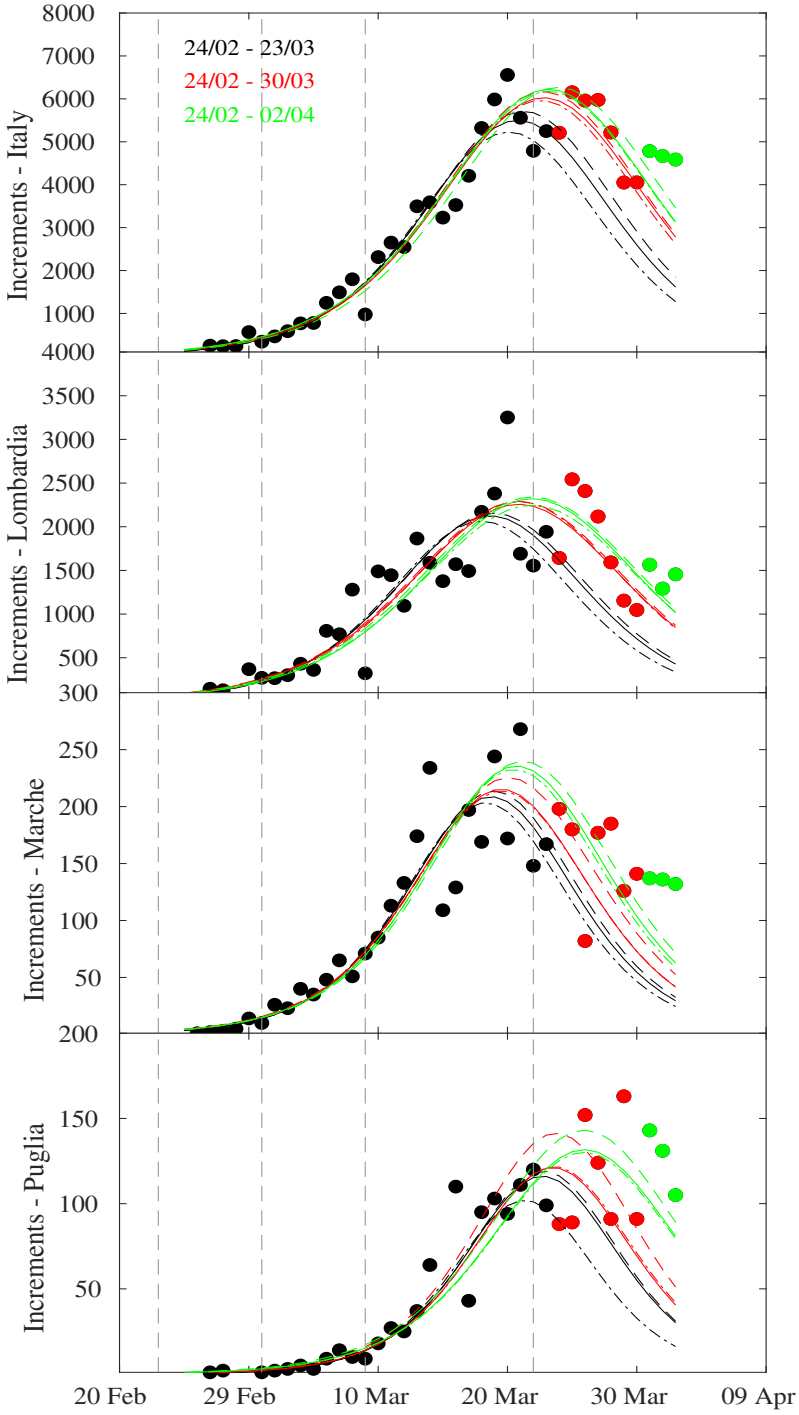

Figure 6: Estimation of daily infections and their peak time during three different time intervals of epidemic across Italian regions, together with the confidence lines. From top to bottom: Italy and three regions (Lombardia, Marche, Puglia). The vertical dashed lines mark the times when Italian government applied lock-down restrictions on 23 February, 01 March, 09 March, and 22 March, respectively.