



Publication Year	2020
Acceptance in OA	2022-02-15T15:31:35Z
Title	On the uncertainty of real-time predictions of epidemic growths: a COVID-19 case study for China and Italy
Authors	ALBERTI, TOMMASO, Davide Faranda
Publisher's version (DOI)	10.1016/j.cnsns.2020.105372
Handle	http://hdl.handle.net/20.500.12386/31393
Journal	COMMUNICATIONS IN NONLINEAR SCIENCE & NUMERICAL SIMULATION
Volume	90

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 Supérieure, 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

38 on asymptotic extrapolation of infection counts obtained on the basis of com-
39 partmental epidemic models as the Susceptible-Exposed-Infected-Recovered
40 (SEIR) model and their variants [25] or on agent-based models [26]. Unfortu-
41 nately, predictions made using these models are extremely sensitive to the un-
42 derlying parameters and the quality of their extrapolation is deeply affected
43 from both the lack of high-quality datasets as well as from the intrinsic sen-
44 sitivity of the dynamics to initial conditions in the growing phase [27]. More-
45 over, in order to provide reliable estimates of asymptotic infection counts, a
46 knowledge of asymptomatic populations is needed. These data are currently
47 almost unavailable and affected by great uncertainties.

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

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

88 2. Data

89 Data for the Chinese provinces are obtained from the data repository for
90 the 2019 Novel Coronavirus Visual Dashboard operated by the Johns Hop-
91 kins University Center for Systems Science and Engineering (JHU CSSE),
92 freely available at <https://github.com/CSSEGISandData/COVID-19>. Fig. 1
93 reports the total number of confirmed infections (left panel), thus including
94 actual positive people to COVID-19, recovered and deaths for China and
95 three Chinese provinces of Beijing, Hubei, and Yunnan, and the daily infec-
96 tions (right panel), during the period between 22 January and 30 March.
97

98 Data for the Italian regions are instead derived from the repository freely
99 available at <https://github.com/pcm-dpc/COVID-19> where data are col-
100 lected from the Italian Protezione Civile from 24 February 2020. Data used
101 here were last downloaded on 02 April, thus covering the period 24 February
102 - 02 April, as shown in Fig. 2. It is evident that although the increments of
103 infections started about 1 month after the Chinese epidemic Italy has fast
104 reached and exceeded the Chinese peak values of ~ 80000 infections. More-
105 over, it is also apparent that epidemic diffusion in China reached its peak
106 within ~ 20 days from the first restriction operated to the Hubei region on 23
107 January. Conversely, the Italian restrictions seem to become more efficient
108 only when the Italian government adopted a lock-down confinement on 9
109 March [17].

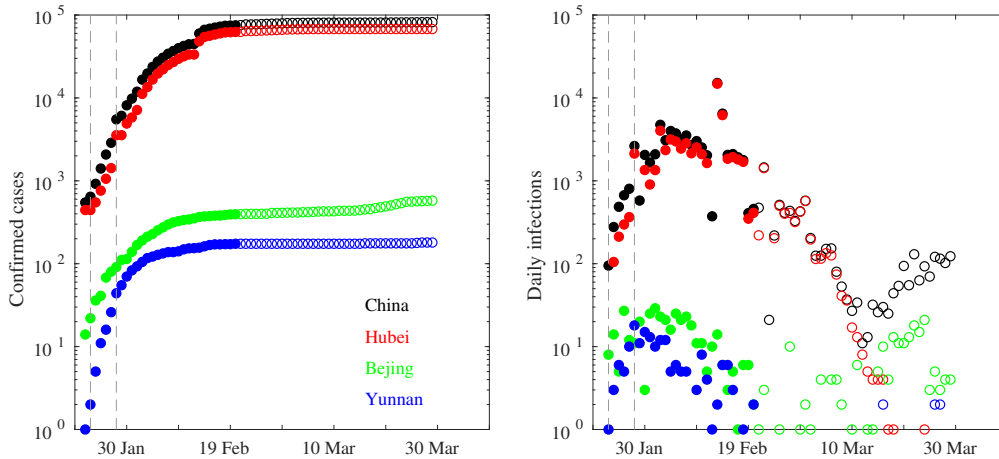


Figure 1: The total number of confirmed infections (left panel) and the daily infections (right panel) for China and three Chinese provinces of Beijing, Hubei, and Yunnan. Filled circles refer to the first 30 days of the epidemic diffusion. The vertical dashed lines mark the times when Chinese government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

110 3. Methods

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

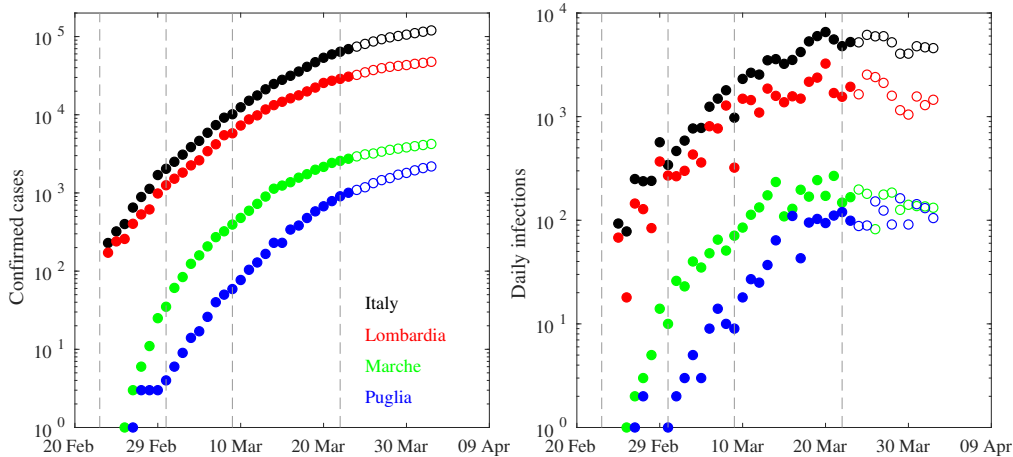


Figure 2: The total number of confirmed infections (left panel) and the daily infections (right panel) for Italy and three Italian regions of Lombardia, Marche, and Puglia. Filled circles refer to the first 30 days of the epidemic diffusion. The vertical dashed lines mark the times when the Italian government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

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

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

133 being α , β , and γ the parameters of the model. They can be fitted, e.g., using
 134 Nonlinear least-squares solver, with the Levenberg-Marquardt algorithm and
 135 the bisquare weight methods to minimize a weighted sum of squares. Here
 136 we use a MATLAB function to perform the fits. As recently pointed out in
 137 [27] in the early stages of the epidemics, the smoothness of COVID-19 cumu-
 138 lative infections data could lead to very uncertain predictions although with
 139 very good R^2 . To avoid this, here we focus only on Chinese and Italian data,
 140 that, to date, represent a mature stage of the epidemics. This implies, as we
 141 will show, that the significance of the logistic fit can be assigned with greater

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

- 153 • the first 30 days of epidemic growth, as reported in Figs. 3-4 by black
154 lines, thus to consider how restrictions measure globally affect the dif-
155 fusion;
- 156 • the first 7 days, roughly corresponding to the time interval during which
157 first restriction measures are adopted both in China and Italy, although
158 not still completely efficient (red lines in Figs. 3-4);
- 159 • the first 14 days, corresponding to the time interval in which the initial
160 confinement measures should lead the first effects (blue lines in Figs. 3-
161 4);
- 162 • the time interval between the 8th and the 14th day to investigate how
163 the epidemic would be grown if starting from initial restrictions (green
164 lines in Figs. 3-4);
- 165 • the time interval between the 15th and the 30th day to investigate the
166 efficiency of restriction measures (magenta lines in Figs. 3-4).

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

176 the efficiency in delivering reliable forecasts at different stages of the epidemic
 177 growth. The test is based on evaluating the maximum distance between the
 178 empirical distribution functions coming from two different samples $x_{1,n}$ and
 179 $x_{2,m}$, being n and m the length of samples. By defining the Kolmogorov-
 180 Smirnov statistic as

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

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

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

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

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

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 .

185
 186 The closer the observed statistics $D_{n,obs}$ is to 0 the more likely it is that
 187 the two samples were drawn from the same distribution with being $D_{n,obs} <$
 188 $D_n(\alpha)$. The use of the K-S test has two main advantages: i) the distribution
 189 of the K-S test statistic itself does not depend on the underlying cumulative
 190 distribution function being tested, and ii) it is an exact test [41, 42, 43, 44].
 191 Moreover, it is specifically designed for testing if data come from a normal,
 192 a log-normal, a Weibull, an exponential, or a logistic distribution [42, 45].
 193 Thus, it is particularly suitable for our investigations, being also a non-
 194 parametric and robust technique since it is not based on strong distributional
 195 assumptions [42, 45, 46, 44].

196 4. Epidemic diffusion through Chinese provinces

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

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

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

Time interval	$D_{n,obs}$			
	China	Hubei	Beijing	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.

237 be released at the early stages of the epidemic diffusion.

238 5. Epidemic diffusion through Italian regions

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

252 5.1. Epidemics growth in Lombardia

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

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

275 5.2. *Epidemics growth in Marche*

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

294 *5.3. Epidemics growth in Puglia*

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

Time interval	$D_{n,obs}$			
	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.

307 It is interesting to note that, although lower values of $D_{n,obs}$ are observed
 308 when a more mature stage of the epidemic growth is considered in the fitting
 309 range, as for example for time intervals from 24 February to 23 March as well
 310 as from 08 to 23 March, the observed values $D_{n,obs}$ are all above the statistical
 311 threshold of $D_{n,th} = 0.3037$. This suggests that a mature stage is, to the date
 312 of 23 March, not yet reached, although Northern and Central regions are
 313 characterized by lower values than the Southern ones, thus possibly related
 314 to the time delayed propagation of epidemic throughout Southern regions.
 315

316 **6. Estimation of infections for Italy and their peak time**

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

Time interval	$D_{n,obs}$			
	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.

324 It is interesting to note that all regions and Italy are characterized by
 325 lower values of $D_{n,obs}$, below the theoretical value $D_{n,th} = 0.3037$, when
 326 including the next 7 days (e.g., by considering the period between 24 Febru-
 327 ary and 30 March) to the logistic fits and when considering the whole time
 328 range (e.g., 24/02 - 02/04). Lombardia presents lower values of the K-S
 329 statistics $D_{n,obs}$ than those for Marche and Puglia, together with a nar-
 330 rower confidence interval when including the successive days, not observed
 331 for both Marche and Puglia. Particularly for Puglia the confidence interval
 332 remains practically unchanged, thus suggesting that logistic fits are not still
 333 stable, possibly due to the fact that Southern regions have not yet reached
 334 a mature stage of the epidemic growth. This difference in terms of sta-
 335 bility of logistic fits as well as on confidence of reliable estimates can be
 336 clearly seen by looking at the behavior of estimated daily increments. Days
 337 of peak significantly depends on the fitting range for Puglia, while the esti-
 338 mation of this quantity is more stable for Lombardia and Marche, as shown
 339 in Fig. 6. Indeed a wider discrepancy is found between daily increments
 340

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

359 7. Conclusion

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

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

378 The most interesting pattern in the time-evolution of the dis-
379 tribution is the observed change from an exponential-like behavior
380 observed at the beginning of the epidemic growth to a sigmoid-like
381 one when first restriction measures are introduced, particularly
382 evident for the Italian case study. Indeed, by evaluating the ex-
383 pected final number of total infections as predicted from logistic fits
384 during the different stages we highlight that reliable estimates can-
385 not be released until more mature stages of the epidemic growth
386 are reached. We show that by only means of the first 7 days,
387 corresponding to the time interval during which first restriction
388 measures are adopted both in China and Italy, an overestimation
389 of the final number of infections of $\sim 65\%$ for China and $\sim 2000\%$
390 for Italy is observed. Conversely, by considering the first 14 days,
391 corresponding to the time interval in which the initial confinement
392 measures should lead the first effects, an underestimation of $\sim 48\%$
393 for China and $\sim 76\%$ for Italy is obtained. A lower underestimation
394 ($\sim 32\%$ and $\sim 69\%$ for China and Italy, respectively) is found when
395 considering the time interval between the 8th and the 14th day, e.g.,
396 by investigating how the epidemic would be grown if starting from
397 initial restrictions only, while a better agreement is found when
398 considering the time interval between the 15th and the 30th day,
399 corresponding to investigate the efficiency of restriction measures,
400 with reduced underestimation of the final number of infections of
401 $\sim 17\%$ for China and $\sim 12\%$ for Italy. Finally, by monitoring the
402 stability of logistic fits as well as their suitability on predicting
403 the number of infections of the successive days (i.e., after the first
404 30 days) we highlight how the uncertainty evolution can be used
405 to track how the epidemics diffused at a regional level, allowing
406 an estimation of the delay in the spread of the virus. Indeed,
407 we found that the uncertainty significantly increases when mov-
408 ing from Northern to Southern regions, where the error is almost
409 comparable with the predicted value, suggesting that, to date, the
410 epidemic peak has not been likely reached for Southern regions,
411 being delayed with respect to Northern and Central ones.

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

415 field. They are crucial to plan the duration of confinement measures and to
416 define the needs for health-care facilities. The aim of this letter was to show
417 that those extrapolations crucially depend not only on the quality of data,
418 but also on the stage of the epidemics, due to the intrinsically non-linear
419 nature of the underlying dynamics. This prevents from performing success-
420 ful long-term extrapolations of the infection counts with statistical models.
421 **As a guideline it is surely helpful to perform logistic fits every**
422 **day and to evaluate the reliability on predicting the next day, and**
423 **then perform a new logistic fit to investigate how the uncertainty**
424 **grown/reduced.** Moreover, reliable estimates are surely affected
425 by possible source of errors in counting infections, thus we suggest
426 to assess the significance of fits to the last data point of the fit-
427 ting range by assuming it could be affected by a $\pm 30\%$ error. This
428 allows us to provide a simple way to estimate confidence intervals
429 [27]. Furthermore, we also suggest not only to exclude the last data
430 point and check fits stability but also to consider to exclude initial
431 point(s) to evaluate how epidemic would be grown if starting from
432 initial restriction measures or how delayed propagation could be
433 present at a regional level.

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

453 **Acknowledgments**

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

456 **References**

- 457 [1] R. F. Service, Coronavirus epidemic snarls science worldwide, *Science*
458 367 (2020) 836–837. doi:10.1126/science.367.6480.836.
- 459 [2] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan,
460 J. Xu, X. Gu, et al., Clinical features of patients infected with 2019
461 novel coronavirus in Wuhan, China, *The Lancet* 395 (2020) 497–506.
462 doi:10.1016/S0140-6736(20)30183-5.
- 463 [3] W.-j. Guan, Z.-y. Ni, Y. Hu, W.-h. Liang, C.-q. Ou, J.-x. He, L. Liu,
464 H. Shan, C.-l. Lei, D. S. Hui, B. Du, L.-j. Li, G. Zeng, K.-Y. Yuen,
465 R.-c. Chen, C.-l. Tang, T. Wang, P.-y. Chen, J. Xiang, S.-y. Li, J.-
466 l. Wang, Z.-j. Liang, Y.-x. Peng, L. Wei, Y. Liu, Y.-h. Hu, P. Peng,
467 J.-m. Wang, J.-y. Liu, Z. Chen, G. Li, Z.-j. Zheng, S.-q. Qiu, J. Luo,
468 C.-j. Ye, S.-y. Zhu, N.-s. Zhong, Clinical characteristics of coronavirus
469 disease 2019 in china, *New England Journal of Medicine* 0 (0) null.
470 doi:10.1056/NEJMoa2002032.
- 471 [4] W. H. Organization, Pneumonia of unknown cause
472 – china, 2020. URL: [https://www.who.int/csr/don/
473 05-january-2020-pneumonia-of-unkown-cause-china/en/](https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/).
- 474 [5] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung,
475 E. H. Lau, J. Y. Wong, X. Xing, N. Xiang, Y. Wu, C. Li, Q. Chen, D. Li,
476 T. Liu, J. Zhao, M. Liu, W. Tu, C. Chen, L. Jin, R. Yang, Q. Wang,
477 S. Zhou, R. Wang, H. Liu, Y. Luo, Y. Liu, G. Shao, H. Li, Z. Tao,
478 Y. Yang, Z. Deng, B. Liu, Z. Ma, Y. Zhang, G. Shi, T. T. Lam, J. T.
479 Wu, G. F. Gao, B. J. Cowling, B. Yang, G. M. Leung, Z. Feng, Early
480 transmission dynamics in wuhan, china, of novel coronavirus-infected
481 pneumonia, *New England Journal of Medicine* 382 (2020) 1199–1207.
482 doi:10.1056/NEJMoa2001316.
- 483 [6] M. Lipsitch, D. L. Swerdlow, L. Finelli, Defining the epidemiology of
484 covid-19 — studies needed, *New England Journal of Medicine* 382 (2020)
485 1194–1196. doi:10.1056/NEJMp2002125.

- 486 [7] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang,
487 W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F.
488 Gao, W. Tan, A novel coronavirus from patients with pneumonia in
489 china, 2019, *New England Journal of Medicine* 382 (2020) 727–733.
490 doi:10.1056/NEJMoa2001017.
- 491 [8] D. Wang, B. Hu, C. Hu, F. Zhu, X. Liu, J. Zhang, B. Wang,
492 H. Xiang, Z. Cheng, Y. Xiong, Y. Zhao, Y. Li, X. Wang,
493 Z. Peng, Clinical Characteristics of 138 Hospitalized Pa-
494 tients With 2019 Novel Coronavirus-Infected Pneumonia in
495 Wuhan, China, *JAMA* 323 (2020) 1061–1069. URL: <https://doi.org/10.1001/jama.2020.1585>. doi:10.1001/jama.2020.1585.
496 arXiv:https://jamanetwork.com/journals/jama/articlepdf/2761044/jama_wang_2020.c
497
- 498 [9] L. T. Phan, T. V. Nguyen, Q. C. Luong, T. V. Nguyen, H. T.
499 Nguyen, H. Q. Le, T. T. Nguyen, T. M. Cao, Q. D. Pham, Im-
500 portation and human-to-human transmission of a novel coronavirus
501 in vietnam, *New England Journal of Medicine* 382 (2020) 872–874.
502 doi:10.1056/NEJMc2001272, pMID: 31991079.
- 503 [10] W. H. Organization, Coronavirus disease 2019 (covid-19) situ-
504 ation report-10, 2020. URL: [https://www.who.int/emergencies/](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/)
505 [diseases/novel-coronavirus-2019/situation-reports/](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/).
- 506 [11] W. E. Parmet, M. S. Sinha, Covid-19 — the law and limits of quar-
507 antine, *New England Journal of Medicine* 0 (0) null. doi:10.1056/
508 NEJMp2004211.
- 509 [12] R. L. Haffajee, M. M. Mello, Thinking globally, acting locally — the
510 u.s. response to covid-19, *New England Journal of Medicine* 0 (0) null.
511 doi:10.1056/NEJMp2006740.
- 512 [13] S. Flaxman, et al., Estimating the number of infections and the
513 impact of non-pharmaceutical interventions on covid-19 in 11
514 european countries, 2020. URL: [https://www.imperial.ac.uk/](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)
515 [media/imperial-college/medicine/sph/ide/gida-fellowships/](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)
516 [Imperial-College-COVID19-Europe-estimates-and-NPI-impact-30-03-2020.](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)
517 [pdf](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).

- 518 [14] C. Rothe, M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, C. Wall-
519 rauch, T. Zimmer, V. Thiel, C. Janke, W. Guggemos, M. Seil-
520 maier, C. Drosten, P. Vollmar, K. Zwirgmaier, S. Zange, R. Wölfel,
521 M. Hoelscher, Transmission of 2019-ncov infection from an asymp-
522 tomatic contact in germany, *New England Journal of Medicine* 382
523 (2020) 970–971. doi:10.1056/NEJMc2001468.
- 524 [15] S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith,
525 A. S. Azman, N. G. Reich, J. Lessler, The Incubation Period of Coro-
526 navirus Disease 2019 (COVID-19) From Publicly Reported Confirmed
527 Cases: Estimation and Application, *Annals of Internal Medicine* (2020).
528 URL: <https://doi.org/10.7326/M20-0504>. doi:10.7326/M20-0504.
529 arXiv:https://annals.org/acp/content_public/journal/aim/0/aim202005050-m20050
- 530 [16] P. M. of Italy, Disposizioni attuative del decreto-legge 23 febbraio
531 2020, n. 6, recante misure urgenti in materia di contenimento e ges-
532 tione dell'emergenza epidemiologica da covid-19, 2020. URL: <https://www.gazzettaufficiale.it/eli/id/2020/02/23/20A01228/sg>.
- 534 [17] P. M. of Italy, Ulteriori disposizioni attuative del decreto-legge 23 feb-
535 braio 2020, n. 6, recante misure urgenti in materia di contenimento e ges-
536 tione dell'emergenza epidemiologica da covid-19, applicabili sull'intero
537 territorio nazionale, 2020. URL: <https://www.gazzettaufficiale.it/eli/id/2020/03/09/20A01558/sg>.
- 539 [18] P. M. of Italy, Ulteriori disposizioni attuative del decreto-legge 23 feb-
540 braio 2020, n. 6, recante misure urgenti in materia di contenimento e ges-
541 tione dell'emergenza epidemiologica da covid-19, applicabili sull'intero
542 territorio nazionale, 2020. URL: <https://www.gazzettaufficiale.it/eli/id/2020/03/11/20A01605/sg>.
- 544 [19] P. M. of Italy, Ulteriori disposizioni attuative del decreto-legge 23 feb-
545 braio 2020, n. 6, recante misure urgenti in materia di contenimento e ges-
546 tione dell'emergenza epidemiologica da covid-19, applicabili sull'intero
547 territorio nazionale, 2020. URL: <https://www.gazzettaufficiale.it/eli/id/2020/03/22/20A01807/sg>.
- 549 [20] L. Rosenbaum, Facing covid-19 in italy — ethics, logistics, and thera-
550 peutics on the epidemic's front line, *New England Journal of Medicine*
551 0 (0) null. doi:10.1056/NEJMp2005492.

- 552 [21] M. Chinazzi, J. T. Davis, M. Ajelli, C. Gioannini, M. Litvi-
553 nova, S. Merler, A. Pastore y Piontti, K. Mu, L. Rossi, K. Sun,
554 C. Viboud, X. Xiong, H. Yu, M. E. Halloran, I. M. Longini,
555 A. Vespignani, The effect of travel restrictions on the spread
556 of the 2019 novel coronavirus (covid-19) outbreak, Science
557 (2020). URL: [https://science.sciencemag.org/content/early/
558 2020/03/05/science.aba9757](https://science.sciencemag.org/content/early/2020/03/05/science.aba9757). doi:10.1126/science.aba9757.
559 arXiv:<https://science.sciencemag.org/content/early/2020/03/05/science.aba9757>.
- 560 [22] A. S. Fauci, H. C. Lane, R. R. Redfield, Covid-19 — navigating the
561 uncharted, New England Journal of Medicine 382 (2020) 1268–1269.
562 doi:10.1056/NEJMe2002387.
- 563 [23] W. H. Organization, Coronavirus disease 2019 (covid-19)
564 situation report-51, 2020. URL: [https://www.who.int/
565 docs/default-source/coronaviruse/situation-reports/
566 20200311-sitrep-51-covid-19.pdf?sfvrsn=1ba62e57\10](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200311-sitrep-51-covid-19.pdf?sfvrsn=1ba62e57\10).
- 567 [24] A. Spinello, A. Saltalamacchia, A. Magistrato, Is the Rigid-
568 ity of SARS-CoV-2 Spike Receptor-Binding Motif the Hall-
569 mark for Its Enhanced Infectivity and Pathogenesis? (2020).
570 URL: [https://chemrxiv.org/articles/Is_the_Rigidity_of_
571 SARS-CoV-2_Spike_Receptor-Binding_Motif_the_Hallmark_
572 for_Its_Enhanced_Infectivity_and_Pathogenesis_/12091260](https://chemrxiv.org/articles/Is_the_Rigidity_of_SARS-CoV-2_Spike_Receptor-Binding_Motif_the_Hallmark_for_Its_Enhanced_Infectivity_and_Pathogenesis_/12091260).
573 doi:10.26434/chemrxiv.12091260.v1.
- 574 [25] F. Brauer, Compartmental models in epidemiology, Lecture Notes in
575 Mathematics, Springer Verlag, 2008, pp. 19–79.
- 576 [26] S. L. Chang, N. Harding, C. Zachreson, O. M. Cliff, M. Prokopenko,
577 Modelling transmission and control of the covid-19 pandemic in aus-
578 tralia, arXiv preprint arXiv:2003.10218 (2020).
- 579 [27] D. Faranda, I. P. Castillo, O. Hulme, A. Jezequel, J. S. W. Lamb,
580 Y. Sato, E. L. Thompson, Asymptotic estimates of sars-cov-2 infection
581 counts and their sensitivity to stochastic perturbation, Chaos under
582 review (2020).
- 583 [28] T. C. Wilcosky, L. E. Chambless, A comparison of direct adjustment
584 and regression adjustment of epidemiologic measures, Journal of chronic
585 diseases 38 (1985) 849–856.

- 586 [29] M. Batista, Estimation of the final size of the covid-19 epidemic,
587 Preprint.] medRxiv (2020).
- 588 [30] C.-H. Li, C.-C. Tsai, S.-Y. Yang, Analysis of the permanence of an
589 SIR epidemic model with logistic process and distributed time delay,
590 Communications in Nonlinear Science and Numerical Simulations 17
591 (2012) 3696–3707. doi:10.1016/j.cnsns.2012.01.018.
- 592 [31] J. Kumar, K. P. S. S. Hembram, Epidemiological study of novel
593 coronavirus (COVID-19), arXiv e-prints (2020) arXiv:2003.11376.
594 arXiv:2003.11376.
- 595 [32] J. Ma, Estimating epidemic exponential growth rate and basic re-
596 production number, Infectious Disease Modelling 5 (2020) 129 –
597 141. URL: [http://www.sciencedirect.com/science/article/pii/
598 S2468042719300491](http://www.sciencedirect.com/science/article/pii/S2468042719300491). doi:[https://doi.org/10.1016/j.idm.2019.12.
599 009](https://doi.org/10.1016/j.idm.2019.12.009).
- 600 [33] G. Chowell, Fitting dynamic models to epidemic outbreaks with
601 quantified uncertainty: A primer for parameter uncertainty, identi-
602 fiability, and forecasts, Infectious Disease Modelling 2 (2017) 379
603 – 398. URL: [http://www.sciencedirect.com/science/article/pii/
604 S2468042717300234](http://www.sciencedirect.com/science/article/pii/S2468042717300234). doi:[https://doi.org/10.1016/j.idm.2017.08.
605 001](https://doi.org/10.1016/j.idm.2017.08.001).
- 606 [34] I. J. Wellock, G. C. Emmans, I. Kyriazakis, Describing and predicting
607 potential growth in the pig, Animal Science 78 (2004) 379–388. doi:10.
608 1017/S1357729800058781.
- 609 [35] B. Gompertz, On the Nature of the Function Expressive of the Law
610 of Human Mortality, and on a New Mode of Determining the Value of
611 Life Contingencies, Philosophical Transactions of the Royal Society of
612 London Series I 115 (1825) 513–583.
- 613 [36] R. Bürger, G. Chowell, L. Y. Lara-Díaz, Comparative analysis of phe-
614 nomenological growth models applied to epidemic outbreaks, Mathe-
615 matical biosciences and engineering : MBE 16 (2019) 4250–4273. URL:
616 <https://doi.org/10.3934/mbe.2019212>. doi:10.3934/mbe.2019212.
- 617 [37] G. Viceconte, N. Petrosillo, Covid-19 r0: Magic number or co-
618 nondrum?, Infectious Disease Reports 12 (2020). URL: <https://>

- 619 www.pagepress.org/journals/index.php/idr/article/view/8516.
620 [doi:10.4081/idr.2020.8516](https://doi.org/10.4081/idr.2020.8516).
- 621 [38] A. N. Kolmogorov, Sulla determinazione empirica di una legge di dis-
622 tribuzione, *Giornale dell'Istituto Italiano degli Attuari* 4 (1933) 83–91.
- 623 [39] N. Smirnov, Table for estimating the goodness of fit of empirical distri-
624 butions, *Ann. Math. Statist.* 19 (1948) 279–281. URL: [https://doi.](https://doi.org/10.1214/aoms/1177730256)
625 [org/10.1214/aoms/1177730256](https://doi.org/10.1214/aoms/1177730256). doi:10.1214/aoms/1177730256.
- 626 [40] M. A. Stephens, Edf statistics for goodness of fit and some compar-
627 isons, *Journal of the American Statistical Association* 69 (1974) 730–
628 737. doi:10.1080/01621459.1974.10480196.
- 629 [41] T. W. Anderson, D. A. Darling, A test of goodness of fit, *Journal of*
630 *the American Statistical Association* 49 (1954) 765–769. doi:10.1080/
631 01621459.1954.10501232.
- 632 [42] G. Marsaglia, W. W. Tsang, J. Wang, Evaluating kolmogorov's distribu-
633 tion, *Journal of Statistical Software, Articles* 8 (2003) 1–4. URL: [https:](https://www.jstatsoft.org/v008/i18)
634 [//www.jstatsoft.org/v008/i18](https://www.jstatsoft.org/v008/i18). doi:10.18637/jss.v008.i18.
- 635 [43] *Kolmogorov–Smirnov Test*, Springer New York, New York, NY, 2008,
636 pp. 283–287. URL: [https://doi.org/10.1007/978-0-387-32833-1_](https://doi.org/10.1007/978-0-387-32833-1_214)
637 [214](https://doi.org/10.1007/978-0-387-32833-1_214). doi:10.1007/978-0-387-32833-1_214.
- 638 [44] M. Marozzi, Nonparametric simultaneous tests for location and scale
639 testing: A comparison of several methods, *Communications in Statis-*
640 *tics - Simulation and Computation* 42 (2013) 1298–1317. doi:10.1080/
641 03610918.2012.665546.
- 642 [45] F. J. Massey, Distribution table for the deviation between two sample
643 cumulatives, *Ann. Math. Statist.* 23 (1952) 435–441. URL: [https://](https://doi.org/10.1214/aoms/1177729388)
644 doi.org/10.1214/aoms/1177729388. doi:10.1214/aoms/1177729388.
- 645 [46] W. Ghidry, E. Lesaffre, G. Verbeke, A comparison of methods for esti-
646 mating the random effects distribution of a linear mixed model, *Statistical*
647 *Methods in Medical Research* 19 (2010) 575–600.

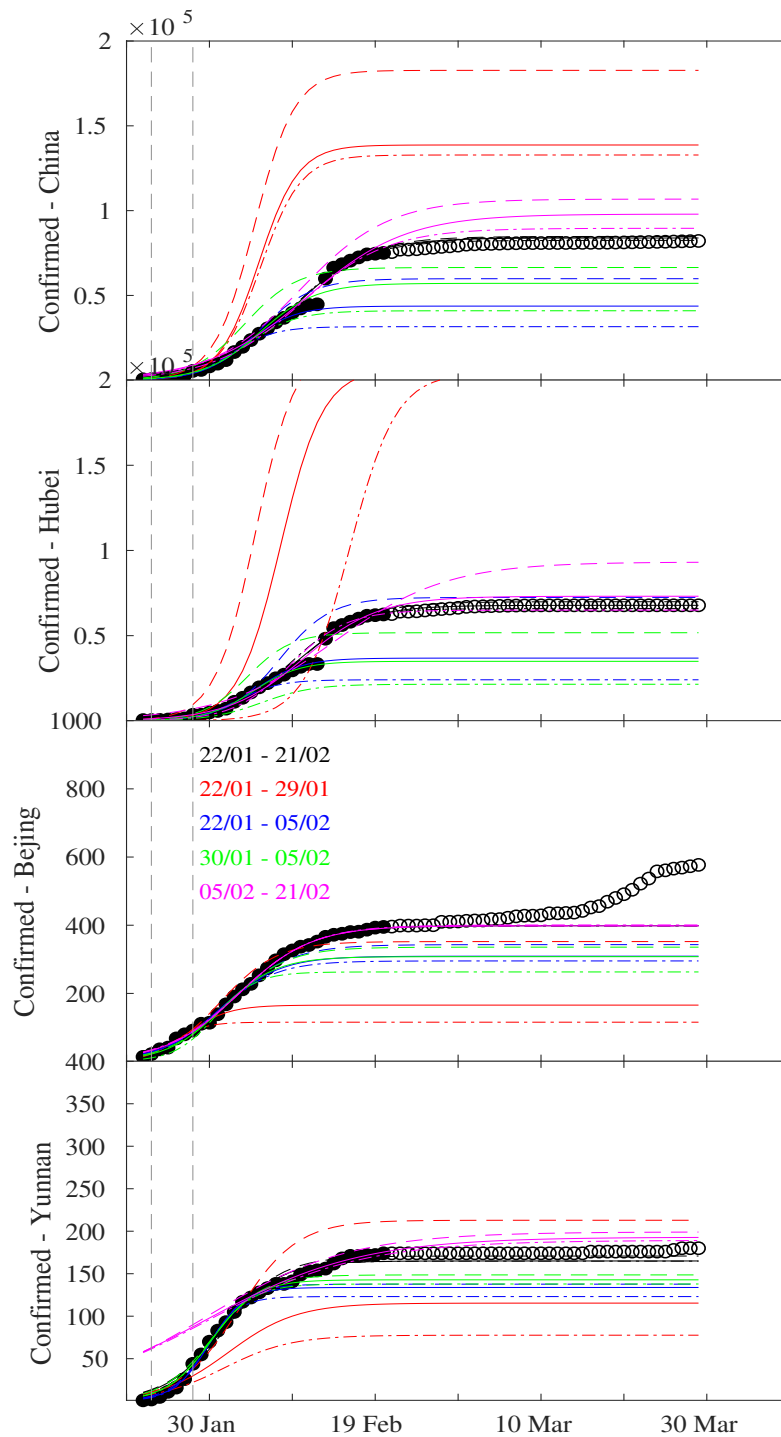


Figure 3: Logistic fits during the different time intervals of epidemic across Chinese provinces, together with the confidence lines. From top to bottom: China and three provinces (Beijing, Hubei, Yunnan). The vertical dashed lines mark the times when Chinese government applied lock-down restrictions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

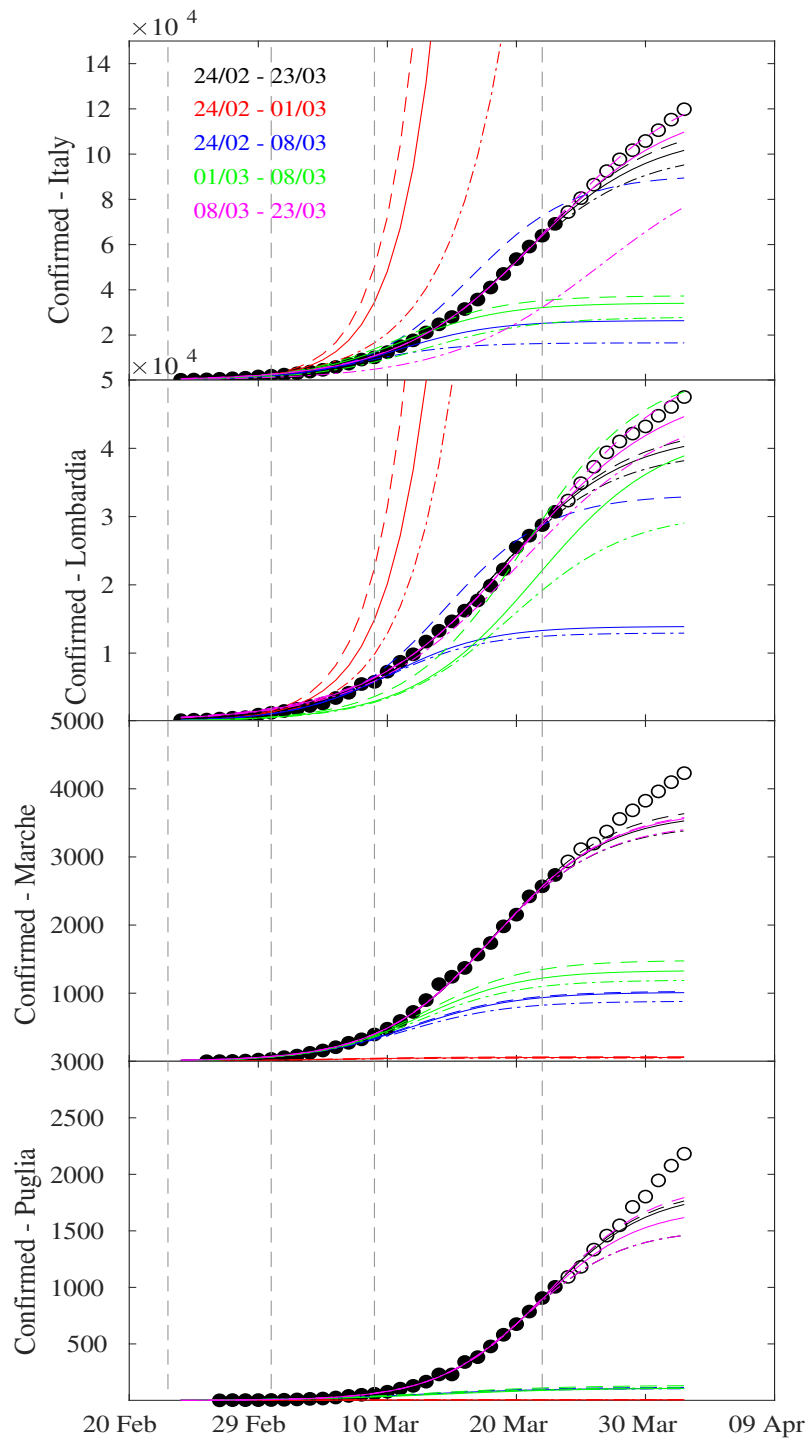


Figure 4: Logistic fits during the different time intervals of epidemic across Italian regions, together with the confidence lines. From top to bottom: Italy and three regions (Lombardia, Marche, Puglia). The vertical dashed lines mark the times when Italian government applied lock-down restrictions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

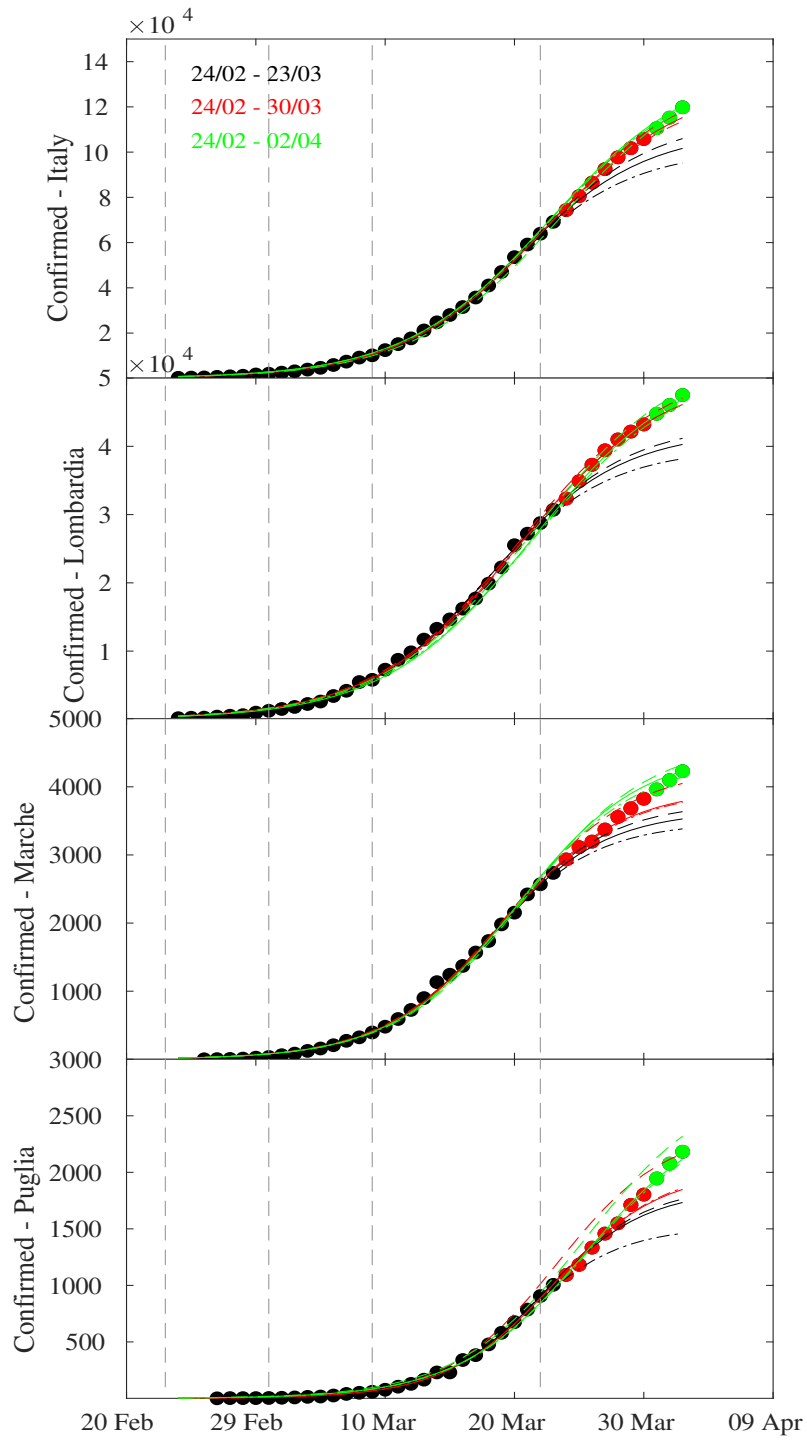


Figure 5: Logistic fits during the different time intervals of epidemic across Italian regions, together with the confidence lines. From top to bottom: Italy and three regions (Lombardia, Marche, Puglia). The vertical dashed lines mark the times when Italian government applied lock-down restrictions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

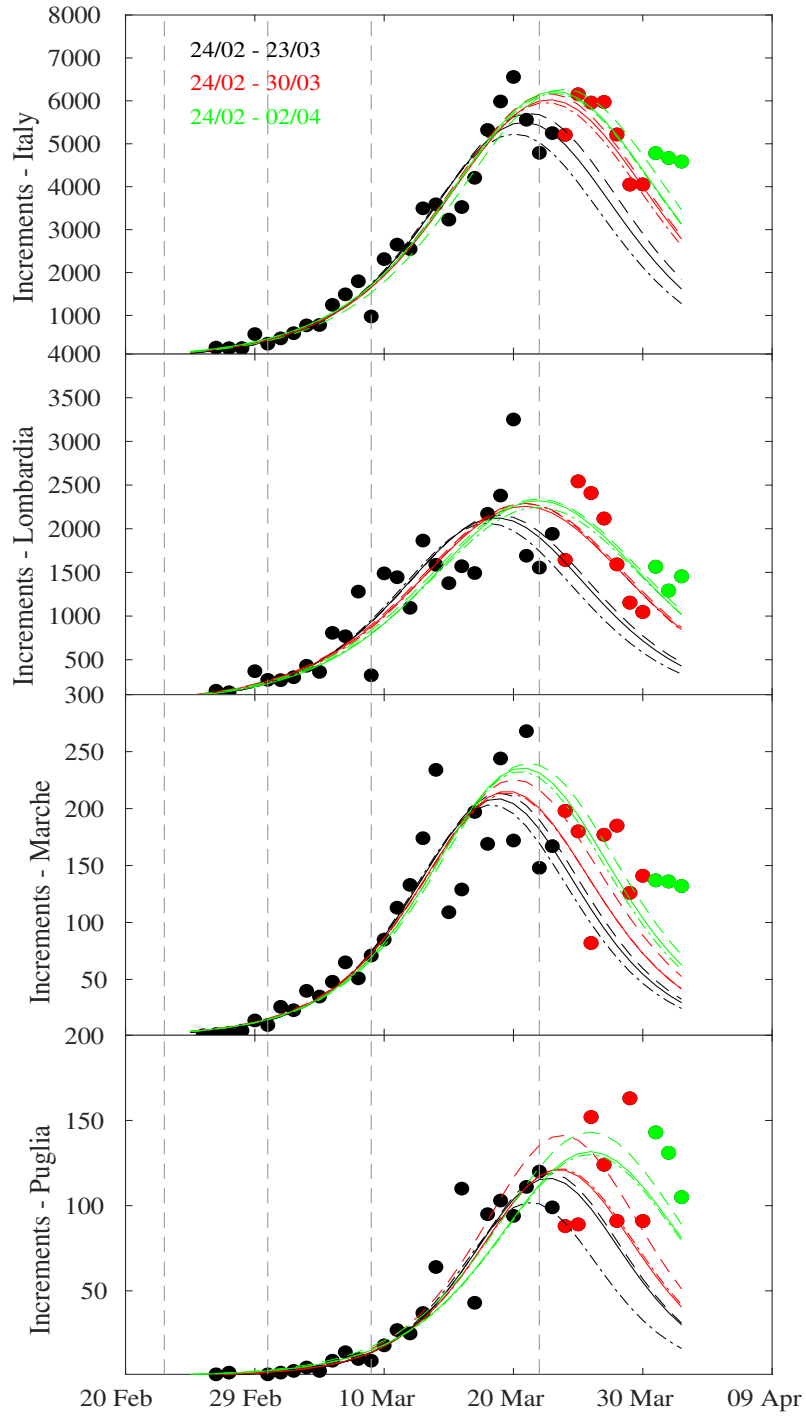


Figure 6: Estimation of daily infections and their peak time during three different time intervals of epidemic across Italian regions, together with the confidence lines. From top to bottom: Italy and three regions (Lombardia, Marche, Puglia). The vertical dashed lines mark the times when Italian government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