COVID-19 Confirmed Cases Prediction as of April 2, 2020

Jointly published by

[1] Chair of Entrepreneurial Risks, D-MTEC, ETH Zurich
[2] Institute of Risk Analysis, Prediction and Management (Risks-X), Academy of Interdisciplinary and Advanced Studies, Southern University of Science and Technology (SUSTech)
[3] Gavekal Intelligence Software

Contacts:
Dr. Ke WU (kwu@ethz.ch) and Prof. Dr. Didier SORNETTE (dsornette@ethz.ch)

Summary of the situation:

-Europe has followed our predicted path yesterday with an 8.4% growth rate today. The inflection point of Europe is still unclear, but we can expect the final number of total confirmed cases to fall in the 1000-2000 range per million population.

-Numbers in Spain, Switzerland, Italy, Austria, and Belgium today continue to support the emerging signs of approaching the inflection point[1], and all of them have converged ensemble distribution of the estimated final total confirmed numbers per million people. Whether Spain and Netherlands would reverse the previous trajectory after the potential inflection point is not clear. The highly uncertain situation in the US continues and we do not expect a quick convergence because the epidemic in the USA is both geographically diverse and at a very early stage. Readers can refer to Supplements to COVID-19 Confirmed Cases Prediction (April 1st, 2020)[2] for our analysis on the US test numbers and the confirmed numbers. Note that the estimated final confirmed numbers tend to underestimate the final results, thus the estimated outbreak progress serves both as a lower bound for future developments and as a guide of the dynamics of the evolution of the epidemics[3].

-Germany and UK have joined France to go back to previous exponential growth today, this makes the future scenario highly uncertain again. Turkey is still at an early stage and will continue its exponential growth in short term. Confirmed infections in Japan continue to follow an exponential growth with a relatively low daily growth rate at 5-10%, while deaths are following a linear trend, indicating that there is not yet a large epidemic in Japan. This could indicate variance in reporting standards of the outbreak world wide.

-The irregular dips and spikes in the data most likely reflect data aggregation and reporting delays where numbers not included one day are included in the following day.

[1] On a logistic curve, the inflection point indicates where the curvature changes its sign. As we model the total number of confirmed cases, it is equal to the peak of the daily increase curve, after which the daily number of cases is decreasing. If the inflection point has been passed, the worst of the outbreak is over.

[2] https://ethz.ch/content/dam/ethz/special-interest/mtec/chair-of-entrepreneurial-risks-dam/documents/Covid-19/Covid_Supplements_1April2020.pdf

[3] One uncertainty with Italy (and other countries) is whether the main outbreak that is focused on the North may spread through other parts of the country. In other words, does the dynamics aggregated over a whole country represent correctly the dynamics in different parts?

Method:

This report updates predictions for the number of COVID-19 confirmed cases at four time horizons (1-day, 5-day, 10-day and end of the outbreak) and for various countries/regions, based on a phenomenological approach detailed in [1]. We employ 3 versions of the generalized logistic growth equation to model the total number of confirmed cases, resulting in a positive, medium and negative scenario for the final expected number of cases. Note that, for countries/regions at early growth stages, the predictions for long-term horizon (10-day and end of the outbreak) are highly uncertain and will vary a lot as the situation changes. The predicted ranges overlap and, as time passes, we anticipate our methodology to zero in on more reliable numbers.

Data source: European Centre for Disease Prevention and Control (ECDC) [2] updated every day at 1pm CET, reflecting data collected up to 6:00 and 10:00 CET. Thus the daily data in some countries is one day delayed compared to other online live sources.

Key Figures & Tables:

-In Table 1, we report the latest confirmed cases per million population and the estimated outbreak progress in the positive and medium scenario (today's confirmed cases divided by the estimated total final confirmed case in positive and now additionally in medium scenarios).

-In Table 2, we report the prediction results in each selected country/region at four time horizons (1-day, 5-day, 10-day and end of the outbreak) in three scenarios. The detailed fitting results for each country/region are plotted in the figures at the end of this report.-

In Figure 1, we present a distribution of the estimated final total confirmed numbers per million population based on the positive and medium scenario.
In Figure 2, we show the 1-day prediction error of yesterday's report.

Comment: We need to emphasize that reported confirmed cases are a leading indicator that is subject to a large number of extraneous variables such as sampling rate[4], sample targeting and reliability of testing. See note at end of this report. The real number of cases in the population is likely to be many multiples higher than those computed from confirmed tests. **We strongly recommend that national governments should publish the number of daily tests and implement random testing (polling) in the population, to facilitate all modeling work and therefore better understanding of the epidemic to help guide appropriate policy responses.**

[4] For instance, The UK is experiencing issues with raising the testing rate linked to a global shortage of certain key reagents and swabs. From since April 1st, all testing is to be targeted at health sector staff and this will obviously bias future data compared with past data.

Table 1. Current confirmed cases per million population and estimated outbreak progress in positive and medium scenarios (today's confirmed cases divided by the estimated total final confirmed cases in positive and medium scenario). Numbers in brackets are 80% confidence intervals. As positive scenarios predict a smaller final number of total infected cases, the outbreak progress is thus larger in the positive scenario. Note that the ranking of this table also reflects intensity of testing.

	Confirmed per Million Population (Mar-31)	Outbreak Progress in Positive Scenario	Outbreak Progress in Medium Scenario
Spain	2186	77.3% (70.7%, 83.2%)	63.4% (55.5%, 70.7%)
Switzerland	2004	74.9% (65.4%, 84.1%)	69.7% (58.0%, 78.9%)
Italy	1830	90.8% (85.4%, 95.4%)	74.9% (70.2%, 79.9%)
Belgium	1223	84.7% (72.1%, 96.5%)	68.0% (54.9%, 76.3%)
Austria	1211	82.6% (75.6%, 89.3%)	81.9% (73.6%, 88.7%)
Germany	887	66.0% (56.6%, 74.5%)	61.3% (50.2%, 70.7%)
France	851	47.0% (24.5%, 62.8%)	4.8% (0.0%, 36.1%)
Netherlands	790	74.3% (68.3%, 79.7%)	68.8% (52.2%, 80.1%)
Europe	667	59.8% (52.0%, 66.9%)	51.5% (43.6%, 59.4%)
United States	662	46.8% (30.2%, 60.3%)	39.6% (19.3%, 54.6%)
United Kingdom	443	24.2% (0.0%, 44.8%)	11.0% (0.0%, 39.0%)
Turkey	190	71.5% (59.2%, 82.5%)	70.1% (53.3%, 81.2%)
Japan	19	11.1% (0.0%, 47.9%)	0.0% (0.0%, 21.9%)

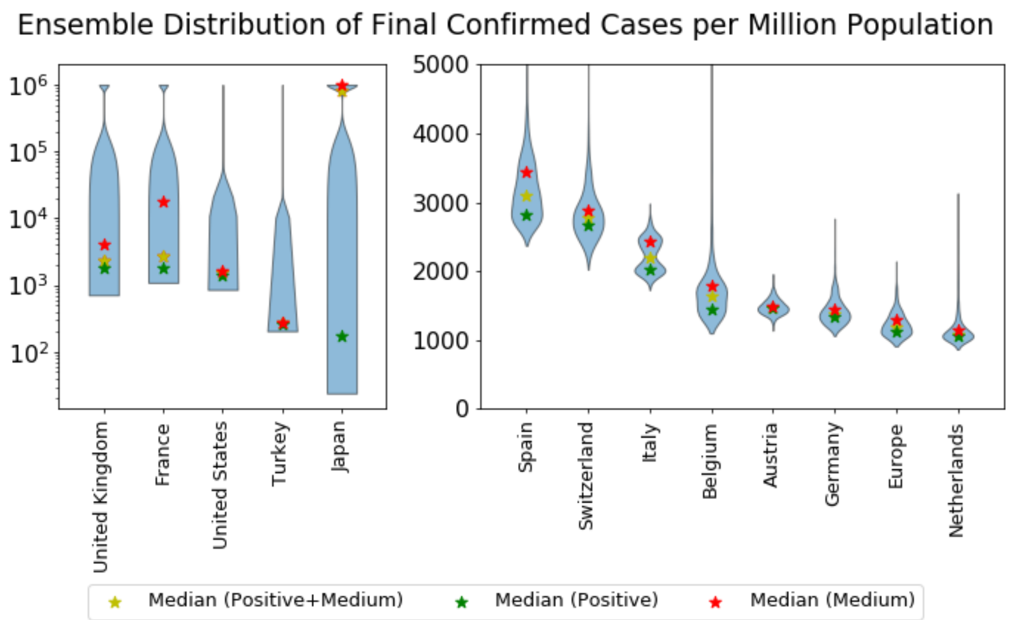


Figure 1. Violin plot of the distributions of the final total number of confirmed cases per million derived by combining the distributions of the positive and medium scenarios. The model setup in the negative scenario does not incorporate a maximum saturation number and thus cannot be used. The yellow star indicates the median prediction for the combined distribution, while the green and red stars indicate the median of the positive and of the medium scenarios respectively. Note that, where we have >1 million infections per 1 million of population, the results are deemed to be unreliable (Table 2).

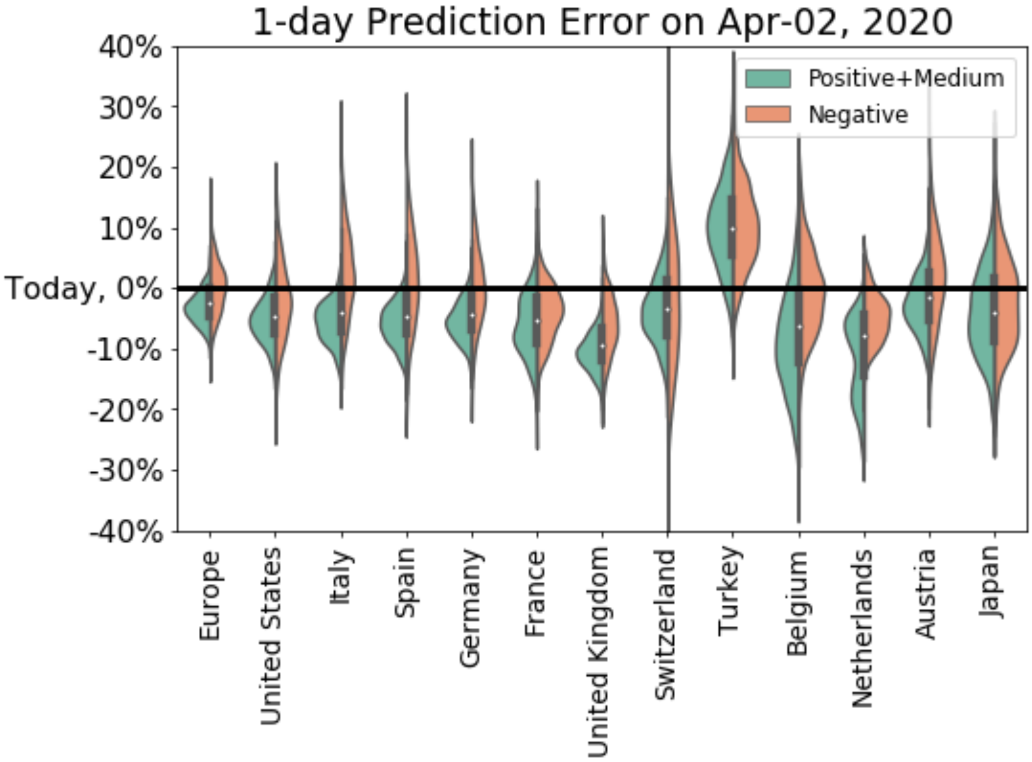


Figure 2. One-day prediction error of the 13 countries/regions. The horizontal line corresponds to today's empirical data. We improve the presentation today compared with previous reports by showing the full distribution (instead of the box plots) in the two scenarios.

Table 2. Predictions for the number of confirmed cases at four time horizons (1-day, 5-day, 10-day and end of the outbreak) and for various countries/regions. The values in parentheses are 80% prediction intervals based on 500 simulations using a negative binomial error structure. In Today's validation column, today's empirical data is presented below yesterday's 1-day predictive interval. "Not reliable" is declared if more than 10% of the simulations produce extreme numbers (larger than total population). All numbers are in thousands.

Country	Scenario*	Today's validation	3-Apr	7-Apr	12-Apr	Final Total Confirmed
Europe	Positive	(466, 505) 498	520 (501, 538)	642 (608, 676)	741 (688, 807)	832 (744, 956)
	Medium	(464, 493) 498	516 (503, 531)	651 (627, 681)	782 (730, 842)	965 (838, 1140)
	Negative	(470, 527) 498	536 (507, 569)	741 (697, 791)	1070 (988, 1170)	Not Reliable
United States	Positive	(195, 218) 217	235 (222, 250)	332 (295, 376)	409 (339, 527)	463 (359, 719)
	Medium	(195, 215) 217	234 (222, 246)	337 (309, 376)	441 (363, 574)	547 (397, 1120)
	Negative	(195, 231) 217	240 (222, 262)	397 (361, 446)	694 (602, 841)	Not Reliable
Italy	Positive	(97.8, 106) 111	106 (101, 110)	114 (109, 120)	119 (114, 125)	122 (116, 129)
	Medium	(103, 110) 111	112 (108, 115)	125 (120, 129)	135 (130, 141)	148 (138, 157)
	Negative	(103, 124) 111	117 (107, 128)	145 (132, 160)	185 (169, 207)	Not Reliable

Spain	Positive	(90.3, 99.3) 102	102 (96.8, 106)	119 (113, 126)	128 (120, 138)	132 (123, 144)
	Medium	(93.4, 101) 102	105 (101, 110)	129 (122, 136)	147 (136, 160)	161 (144, 184)
	Negative	(93.5, 114) 102	110 (99.2, 123)	157 (139, 176)	232 (202, 275)	Not Reliable
Germany	Positive	(66.8, 73) 73.5	77.1 (73.3, 81.2)	92.6 (86.6, 99.2)	103 (94.5, 115)	111 (98.7, 130)
	Medium	(66.4, 72.4) 73.5	76.3 (72.8, 79.9)	93.6 (87.8, 99.6)	107 (97.1, 119)	120 (104, 147)
	Negative	(68, 80.1) 73.5	80.3 (73.8, 86.1)	111 (102, 121)	161 (145, 179)	Not Reliable
France	Positive	(48.8, 56.7) 57	57.7 (54.3, 61.9)	79.4 (70.9, 89.6)	101 (82.4, 131)	121 (90.8, 233)
	Medium	(51.1, 57.6) 57	59.6 (56.3, 63)	87.2 (79.2, 94.2)	133 (105, 151)	Not Reliable
	Negative	(51, 57.9) 57	60.2 (57.1, 63.7)	89.4 (84.2, 95.5)	141 (128, 154)	Not Reliable
United Kingdom	Positive	(24.9, 27.5) 29.5	32 (30.1, 34)	51.1 (44.5, 59.5)	77.4 (56.9, 123)	Not Reliable
	Medium	(25.5, 27.7) 29.5	31.7 (30.3, 33.5)	52.2 (46, 57.4)	88.8 (61.1, 111)	Not Reliable
	Negative	(25.8, 29.6) 29.5	32.1 (30.4, 33.7)	54.6 (51.3, 58.9)	102 (89.6, 116)	Not Reliable
Switzerland	Positive	(16.1, 18.8) 17.1	18.2 (16.9, 19.6)	20.5 (18.7, 22.4)	21.9 (19.8, 24.5)	22.8 (20.3, 26.1)
	Medium	(16.1, 18.3) 17.1	18.2 (17, 19.3)	20.8 (19.3, 22.5)	22.8 (20.7, 25.4)	24.5 (21.6, 29.4)
	Negative	(15.4, 20.6) 17.1	18.9 (16.3, 21.7)	24.7 (21.4, 28.8)	33 (28.1, 40.2)	Not Reliable
Turkey	Positive	(15.9, 18.9) 15.7	16.4 (15.2, 17.7)	20.6 (18.4, 23.4)	21.7 (18.9, 25.9)	21.9 (19, 26.5)
	Medium	(15.7, 18.4) 15.7	16.4 (15.3, 17.7)	20.8 (18.6, 24.2)	22.1 (19.2, 28)	22.4 (19.3, 29.4)
	Negative	(15.9, 18.7) 15.7	17.5 (16, 19.3)	31.4 (27.9, 36.6)	57.5 (48.1, 74.7)	Not Reliable
Belgium	Positive	(10.8, 13.2) 14	13.3 (12.2, 14.6)	15.6 (13.9, 17.6)	16.3 (14.4, 19)	16.5 (14.5, 19.4)
	Medium	(12.2, 14.6) 14	14.5 (13.3, 15.6)	17.8 (16.3, 19.9)	19.7 (17.8, 23.1)	20.5 (18.3, 25.4)
	Negative	(12.6, 15.2) 14	15.1 (13.7, 16.6)	22.5 (20.4, 24.9)	34.9 (30.8, 40.8)	Not Reliable
Netherlands	Positive	(13.1, 14.1) 13.6	14.5 (13.9, 15.1)	16.6 (15.8, 17.6)	17.7 (16.7, 19.1)	18.3 (17.1, 19.9)
	Medium	(11.3, 13.1) 13.6	13.4 (12.6, 14.2)	16.3 (15, 18)	18.3 (16.3, 21.7)	19.8 (17, 26.1)
	Negative	(13.5, 15) 13.6	15.4 (14.5, 16.3)	21.4 (20, 22.8)	30.8 (28.4, 33.5)	Not Reliable
Austria	Positive	(9.82, 11.1) 10.7	11 (10.3, 11.7)	12.1 (11.4, 13)	12.7 (11.8, 13.8)	13 (12, 14.2)
	Medium	(9.77, 11.1) 10.7	10.9 (10.3, 11.6)	12.1 (11.4, 13)	12.8 (11.9, 13.9)	13.1 (12.1, 14.6)

	Negative	(9.87, 11.9) 10.7	11.4 (10.4, 12.7)	14.9 (13.4, 16.6)	20 (17.6, 22.7)	Not Reliable
Japan	Positive	(1.87, 2.25) 2.38	2.23 (2.04, 2.44)	3.15 (2.79, 3.52)	4.7 (3.54, 5.76)	Not Reliable
	Medium	(1.95, 2.34) 2.38	2.38 (2.18, 2.58)	3.3 (2.99, 3.65)	4.96 (4.17, 5.67)	Not Reliable
	Negative	(1.94, 2.35) 2.38	2.37 (2.19, 2.6)	3.31 (3.02, 3.67)	5.03 (4.42, 5.69)	Not Reliable
Iran	Positive	(35.4, 44.5) 47.6	42.6 (38, 46.7)	52.1 (45.1, 62.9)	59.4 (48.3, 85.3)	Not Reliable
	Medium	(43, 49.1) 47.6	49 (46.5, 52.2)	62.7 (58.2, 67.4)	82.4 (72.4, 91.4)	Not Reliable
	Negative	(43.5, 49.1) 47.6	49.6 (46.8, 52.4)	63.9 (60.2, 68)	85.8 (79.4, 91.8)	Not Reliable

* Note:

-The scenarios are based on the final total confirmed numbers. The positive and medium scenarios are derived from the Generalized Logistic Model and the Logistic Model. The model with the lower mean predicted final total confirmed number K, is classified as the positive scenario, and the other one is classified as the medium scenario. The negative scenario is based on the Generalized Growth model, which should only describe the early stage of the epidemic outbreak and is therefore least reliable for countries in the more mature stage.

-Trajectories from Iran have largely deviated from a typical logistic type growth (S curve), and can't be properly described by our models. Although we still report its calibration results in Table 1, they should not be taken as reliable in all scenarios and time horizons. This is probably a result of unreliable reported data from Iran.

Limitations of using the statistics of reported confirmed number

It is important to understand what our prediction models show. The predictions are based on cases identified on the basis of testing and they therefore predict the numbers of future positive tests. Relating positive test results to real levels of infection is subject to a large number of biases. It is a fact that the real number of infections is far higher than those recorded in positive tests since only a tiny fraction of any population has been tested. It is also the case that, in most countries, testing is biased towards those who think they are infected. The first bias, therefore, will underestimate the real number of infections while the second bias will tend to overestimate since it is biased towards those who think they are ill.

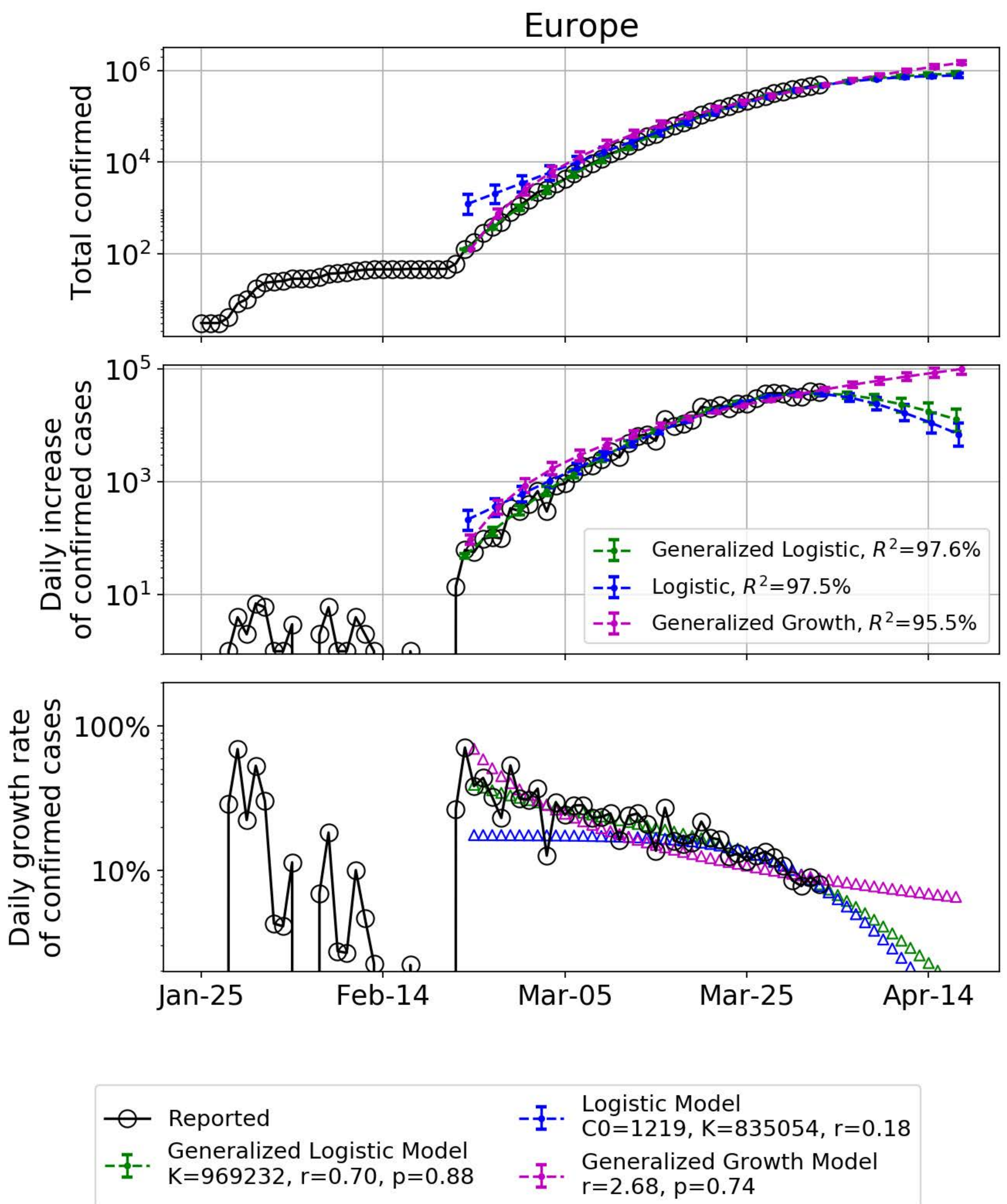
There are further complications. Depending on the testing protocols used, in some instances false positive results have been obtained. In other words, someone without the disease tested positive, probably because they were infected with some other coronavirus. And in other cases, false negative results were obtained, as was the case with the early testing deployed in the USA.

One final complication is the fact that tests are conducted sequentially over time. They do not represent a snapshot of a day in time. Many of those tested early, giving a negative result, may today get a positive result. And many, who tested positive early on, may today be cured.

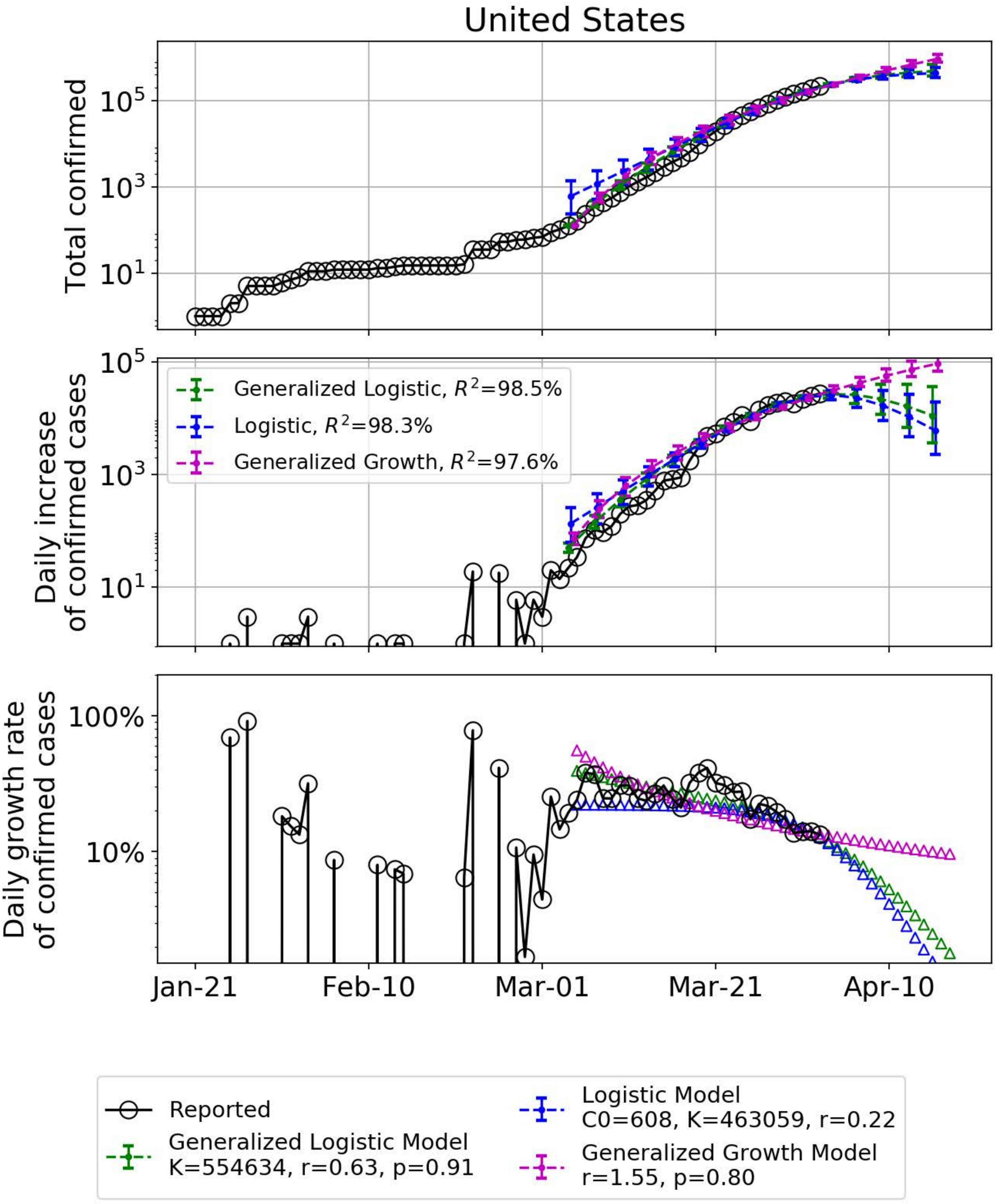
We anticipate that, over time, our methodology will improve and will provide a more accurate picture of the true levels of infection and where they are headed.

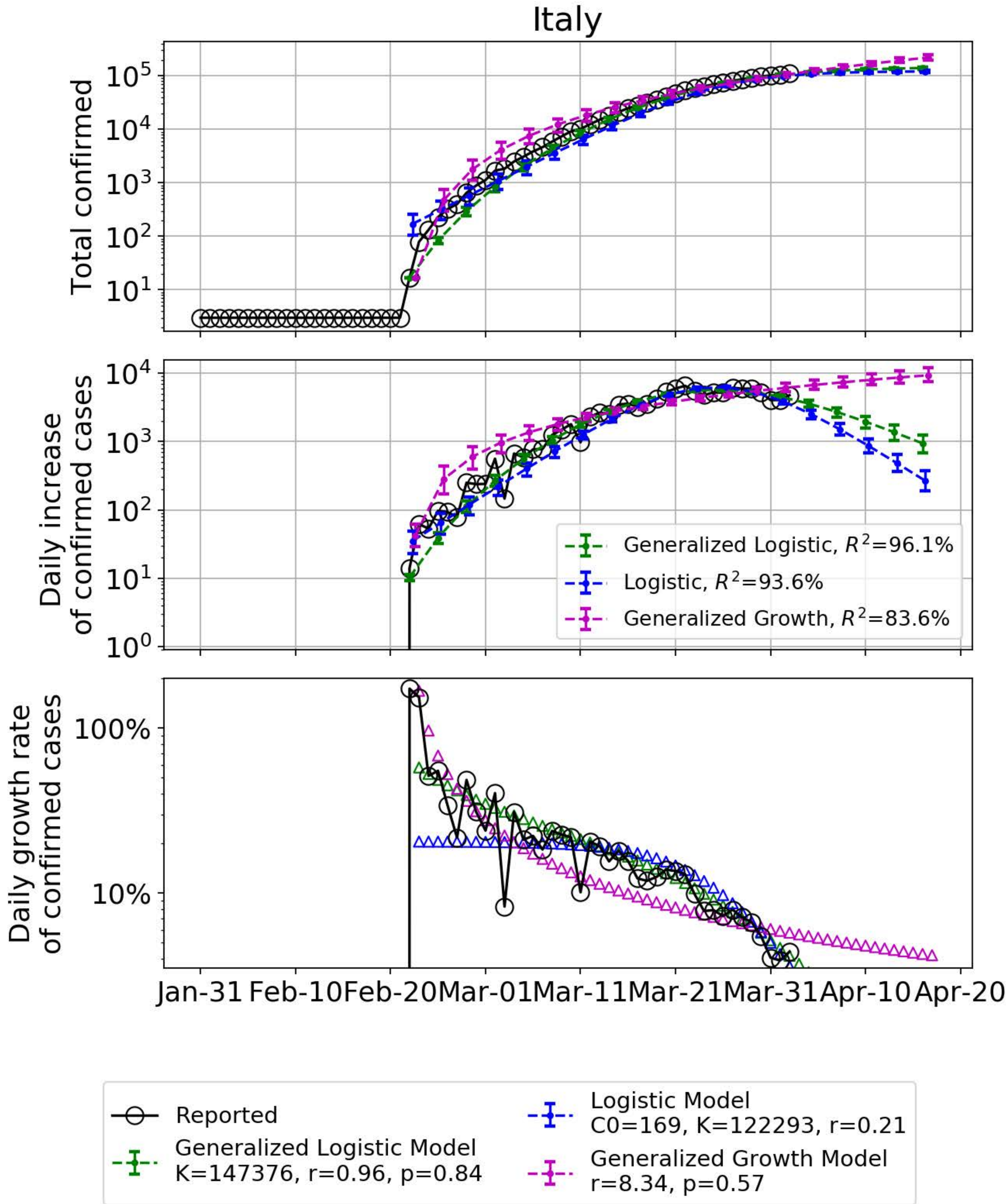
[1] Ke Wu, Didier Darcet, Qian Wang and Didier Sornette, Generalized logistic growth modeling of the COVID-19 outbreak in 29 provinces in China and in the rest of the world, preprint at http://arxiv.org/abs/2003.05681 and
medRxiv: https://medrxiv.org/cgi/content/short/2020.03.11.20034363v1

[2] https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases

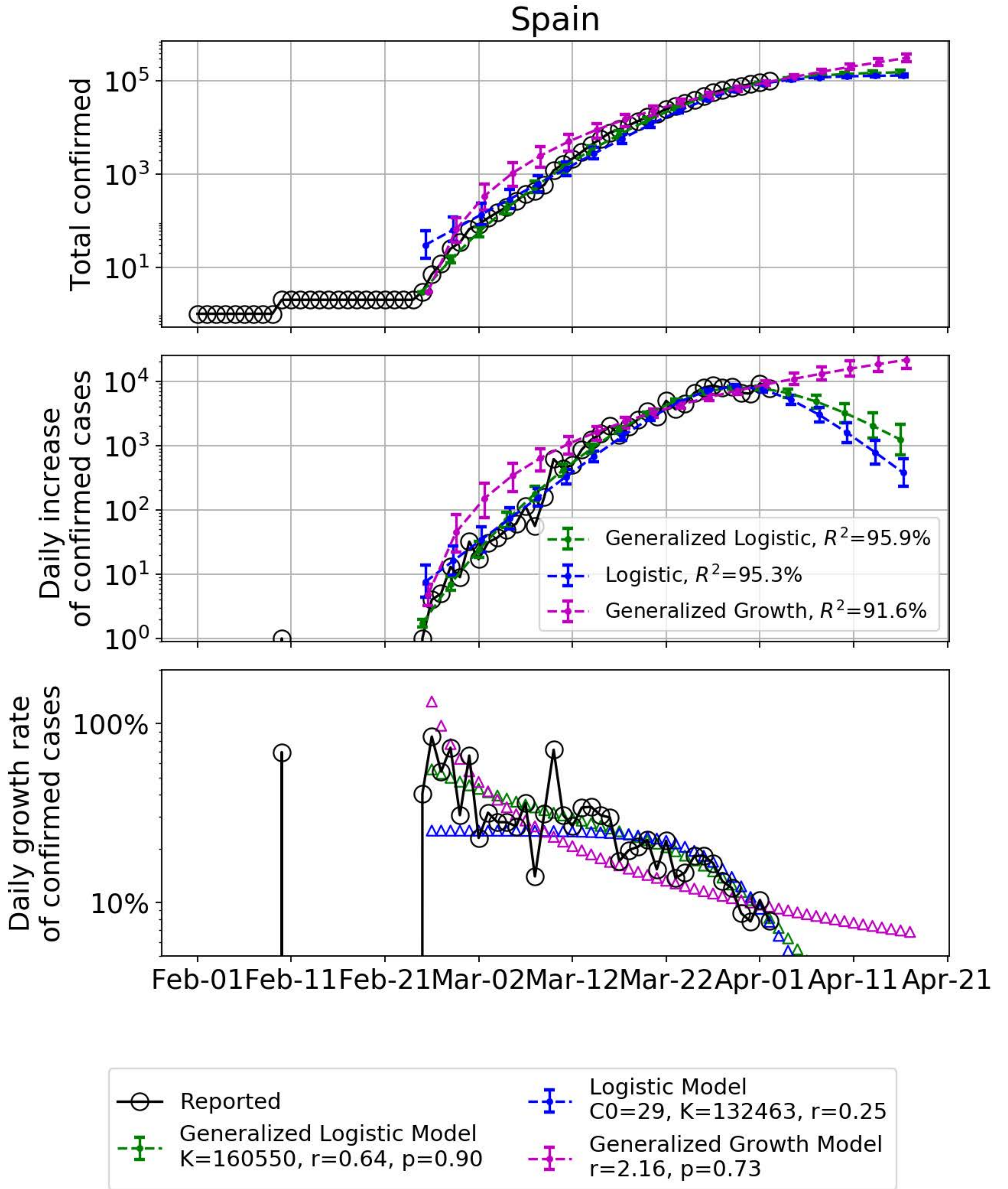


Europe
Total confirmed
Daily increase of confirmed cases
Daily growth rate of confirmed cases
Generalized Logistic, R^2=97.6%
Logistic, R^2=97.5%
Generalized Growth, R^2=95.5%
Jan-25
Feb-14
Mar-05
Mar-25
Apr-14
Reported
Generalized Logistic Model
K=969232, r=0.70, p=0.88
Logistic Model
C0=1219, K=835054, r=0.18
Generalized Growth Model
r=2.68, p=0.74

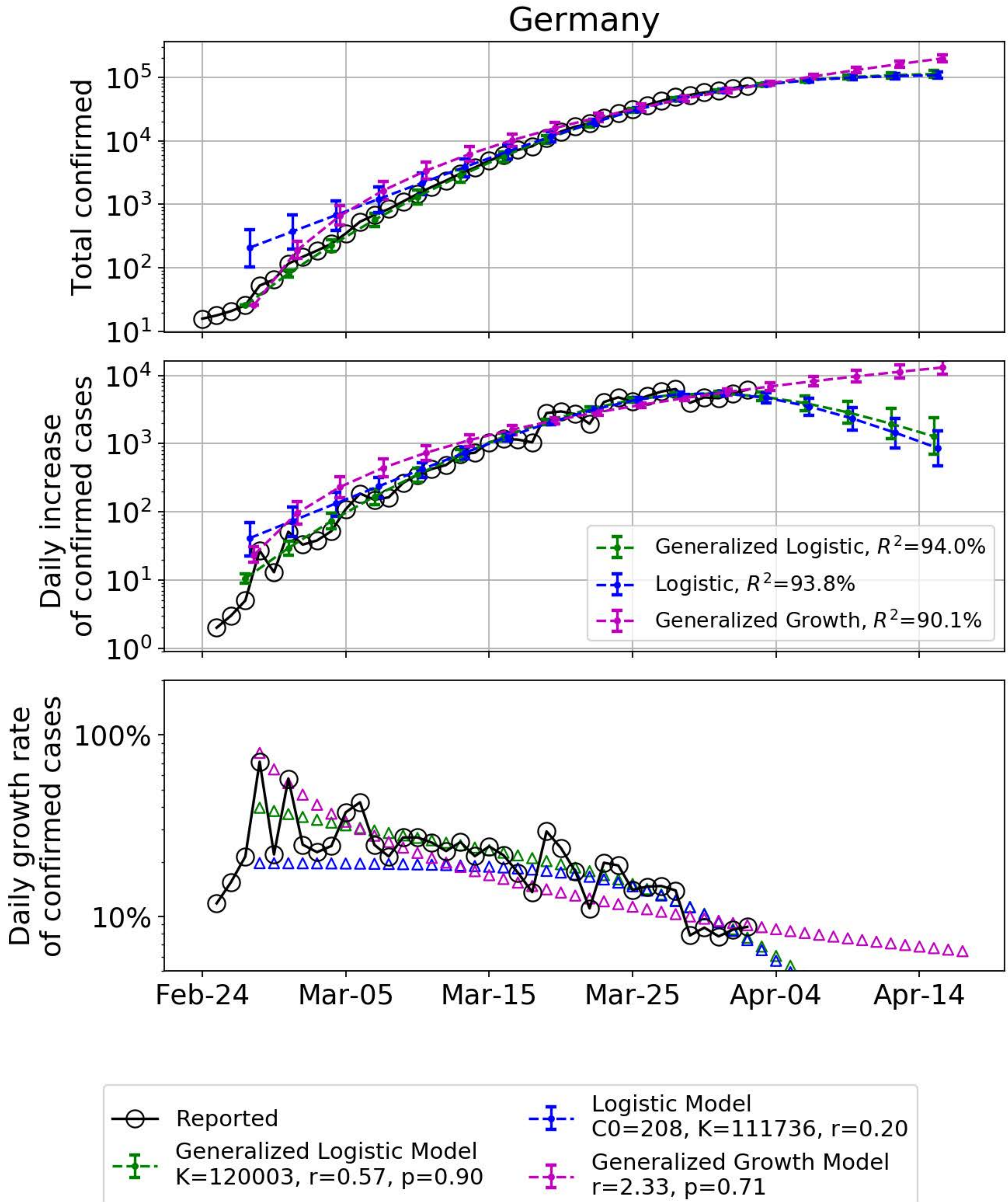

United States
Total confirmed
10^5
10^3
10^1
Daily increase of confirmed cases
Generalized Logistic, R^2=98.5%
Logistic, R^2=98.3%
Generalized Growth, R^2=97.6%
Daily growth rate of confirmed cases
100%
10%
Jan-21
Feb-10
Mar-01
Mar-21
Apr-10
Reported
Generalized Logistic Model
K=554634, r=0.63, p=0.91
Logistic Model
C0=608, K=463059, r=0.22
Generalized Growth Model
r=1.55, p=0.80



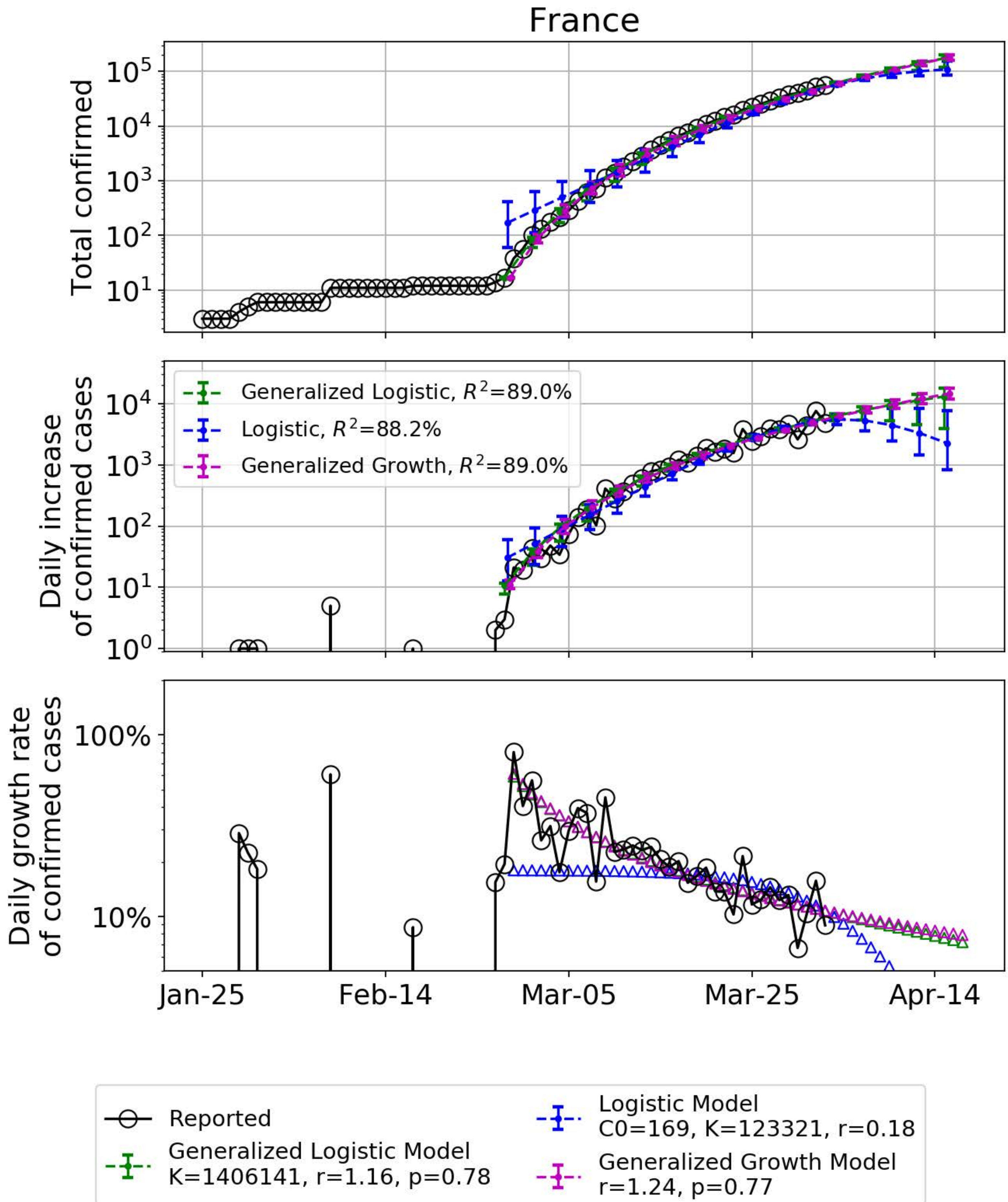
Italy
Total confirmed
Daily increase of confirmed cases
Daily growth rate of confirmed cases
Generalized Logistic, R^2=96.1%
Logistic, R^2=93.6%
Generalized Growth, R^2=83.6%
Jan-31
Feb-10
Feb-20
Mar-01
Mar-11
Mar-21
Mar-31
Apr-10
Apr-20
Reported
Generalized Logistic Model
K=147376, r=0.96, p=0.84
Logistic Model
C0=169, K=122293, r=0.21
Generalized Growth Model
r=8.34, p=0.57

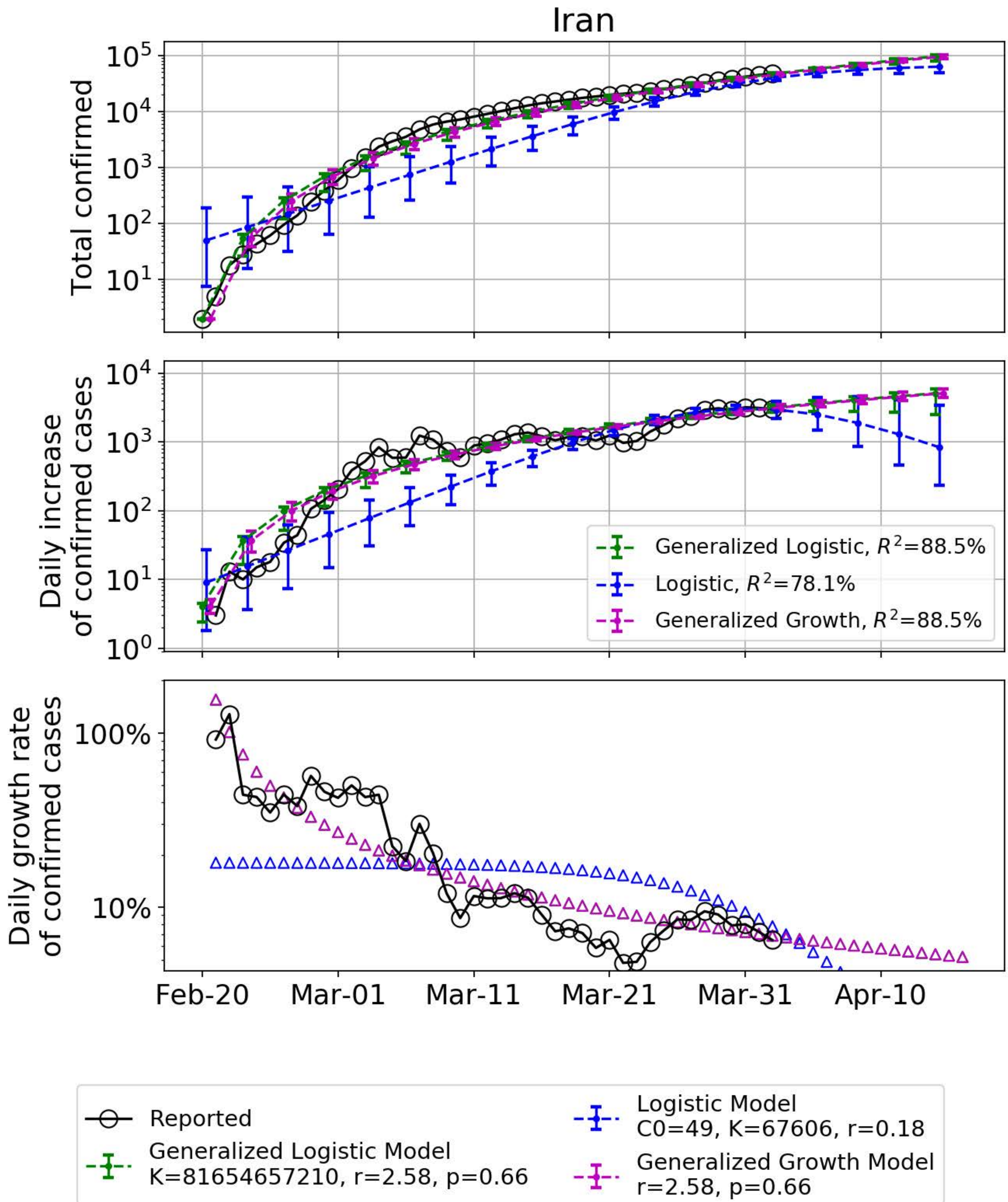


Spain
Total confirmed
Daily increase of confirmed cases
Daily growth rate of confirmed cases
Generalized Logistic, R^2=95.9%
Logistic, R^2=95.3%
Generalized Growth, R^2=91.6%
Feb-01
Feb-11
Feb-21
Mar-02
Mar-12
Mar-22
Apr-01
Apr-11
Apr-21
Reported
Generalized Logistic Model
K=160550, r=0.64, p=0.90
Logistic Model
C0=29, K=132463, r=0.25
Generalized Growth Model
r=2.16, p=0.73

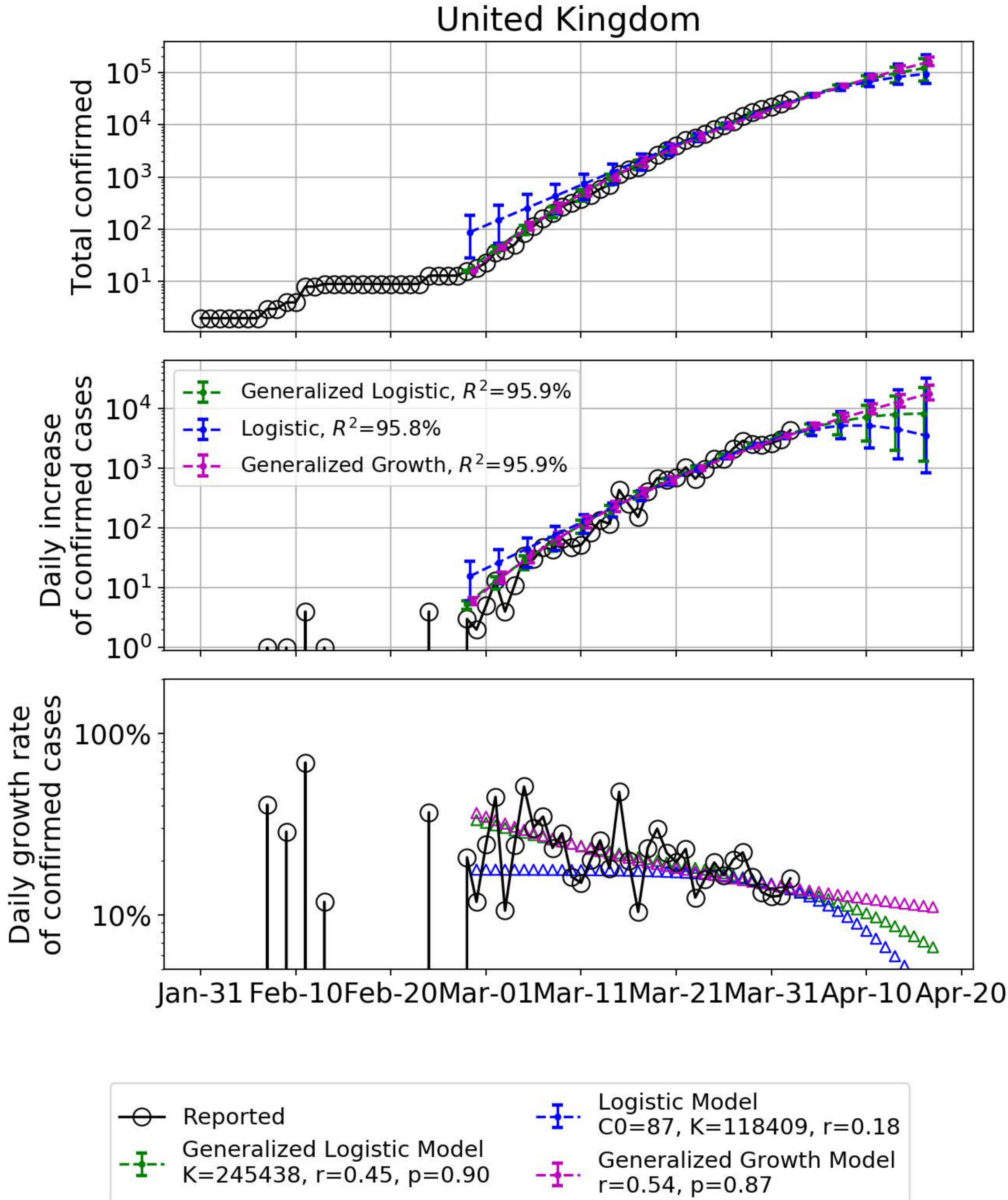


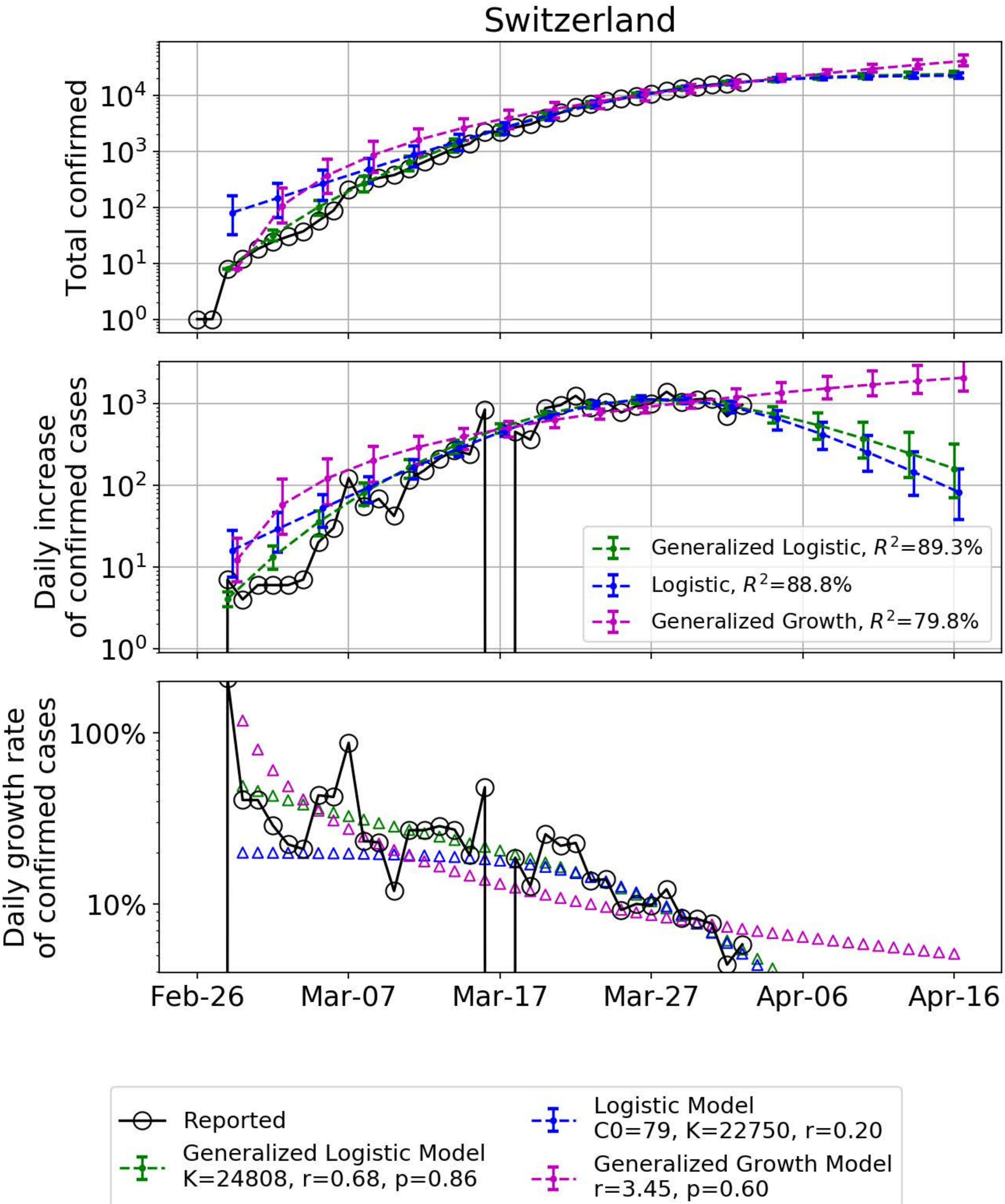
Germany
Total confirmed
Daily increase of confirmed cases
Daily growth rate of confirmed cases
Generalized Logistic, R^2=94.0%
Logistic, R^2=93.8%
Generalized Growth, R^2=90.1%
100%
10%
Feb-24
Mar-05
Mar-15
Mar-25
Apr-04
Apr-14
Reported
Generalized Logistic Model
K=120003, r=0.57, p=0.90
Logistic Model
C0=208, K=111736, r=0.20
Generalized Growth Model
r=2.33, p=0.71


France
Total confirmed
Daily increase of confirmed cases
Daily growth rate of confirmed cases
Generalized Logistic, R^2=89.0%
Logistic, R^2=88.2%
Generalized Growth, R^2=89.0%
Jan-25
Feb-14
Mar-05
Mar-25
Apr-14
Reported
Generalized Logistic Model
K=1406141, r=1.16, p=0.78
Logistic Model
C0=169, K=123321, r=0.18
Generalized Growth Model
r=1.24, p=0.77

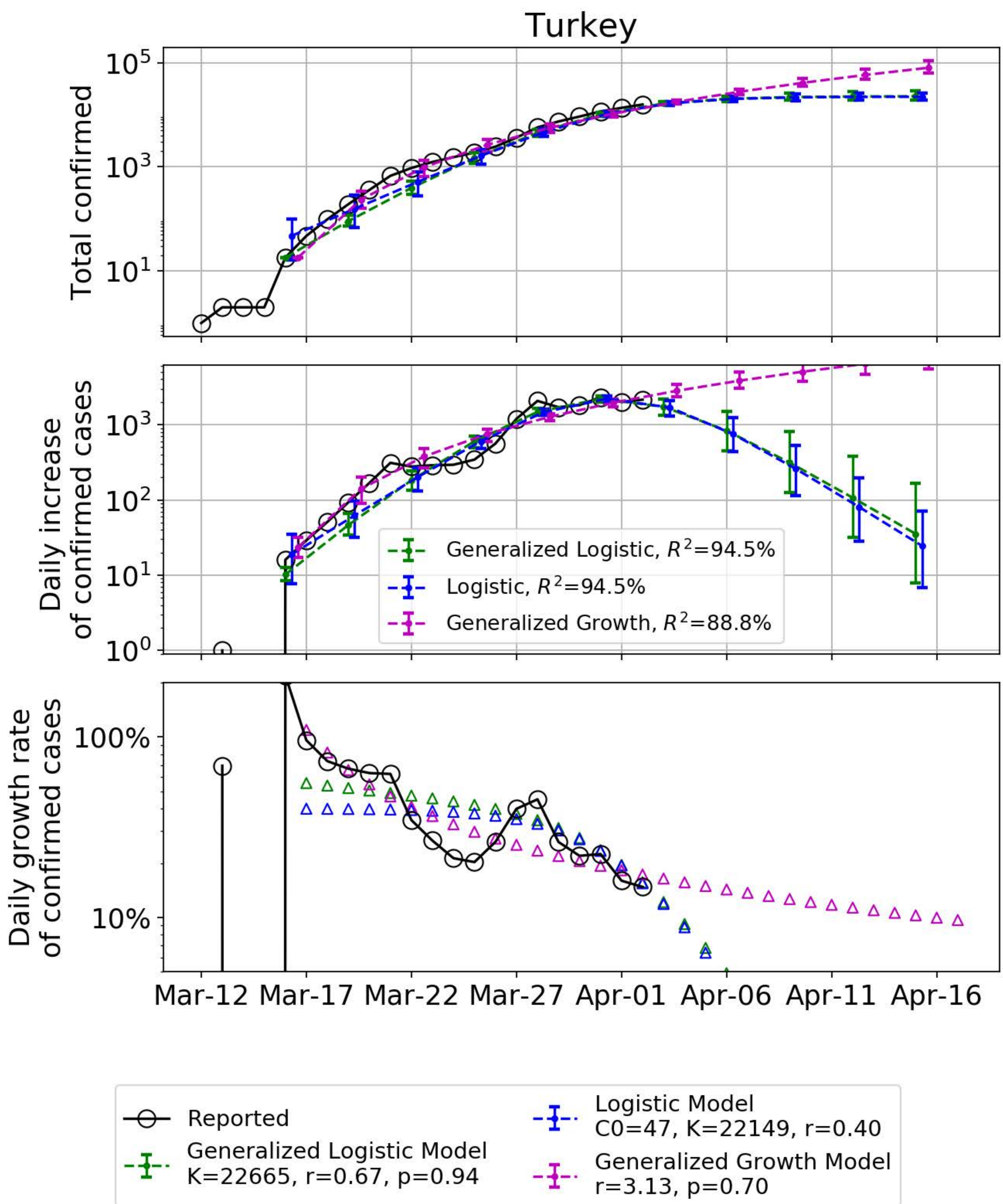


Iran
Total confirmed
Daily increase of confirmed cases
Daily growth rate of confirmed cases
Generalized Logistic, R^2=88.5%
Logistic, R^2=78.1%
Generalized Growth, R^2=88.5%
Feb-20
Mar-01
Mar-11
Mar-21
Mar-31
Apr-10
Reported
Generalized Logistic Model
K=81654657210, r=2.58, p=0.66
Logistic Model
C0=49, K=67606, r=0.18
Generalized Growth Model
r=2.58, p=0.66

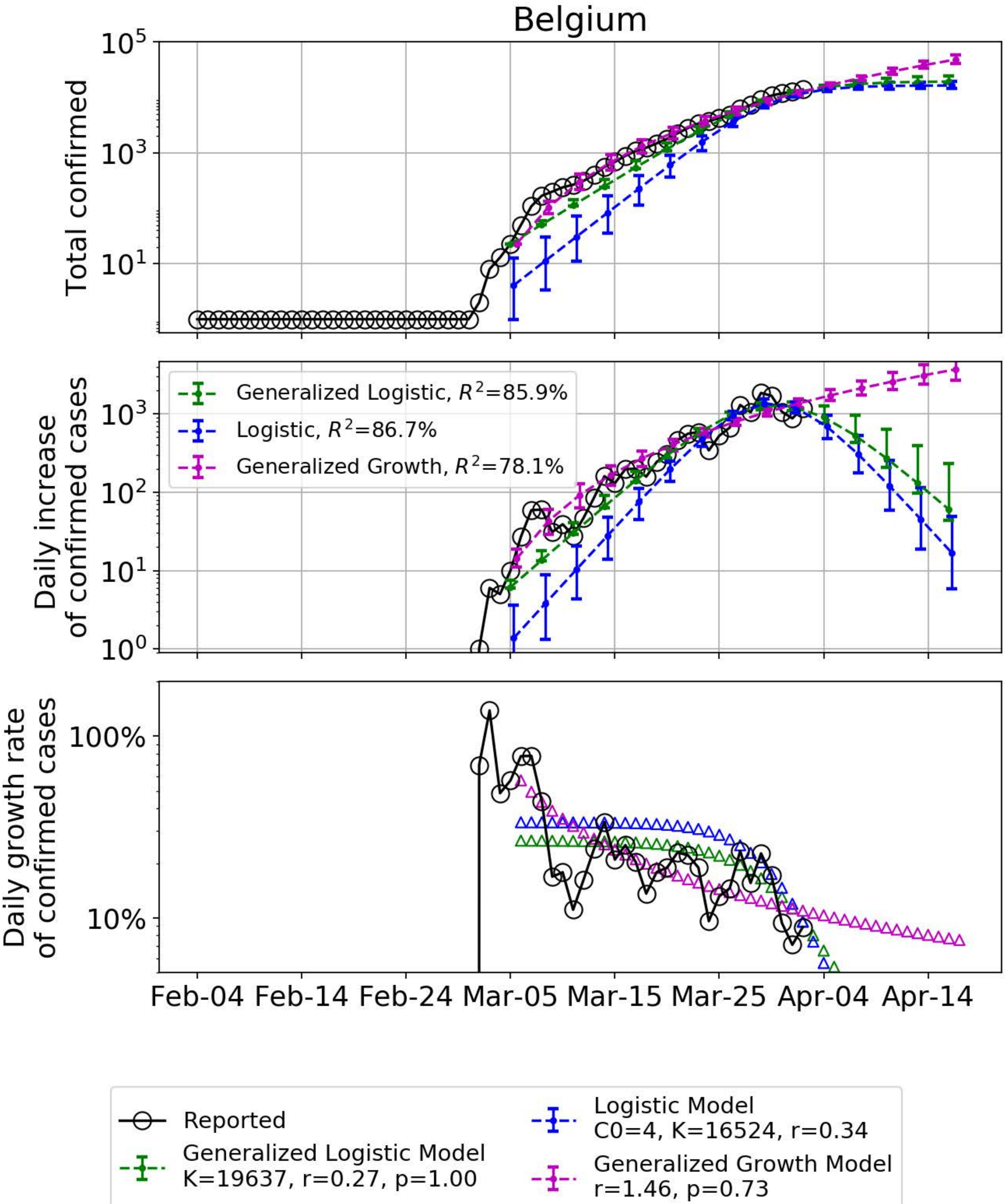

United Kingdom
Total confirmed
Daily increase of confirmed cases
Daily growth rate of confirmed cases
Generalized Logistic, R^2=95.9%
Logistic, R^2=95.8%
Generalized Growth, R^2=95.9%
Jan-31
Feb-10
Feb-20
Mar-01
Mar-11
Mar-21
Mar-31
Apr-10
Apr-20
Reported
Generalized Logistic Model
K=245438, r=0.45, p=0.90
Logistic Model
C0=87, K=118409, r=0.18
Generalized Growth Model
r=0.54, p=0.87



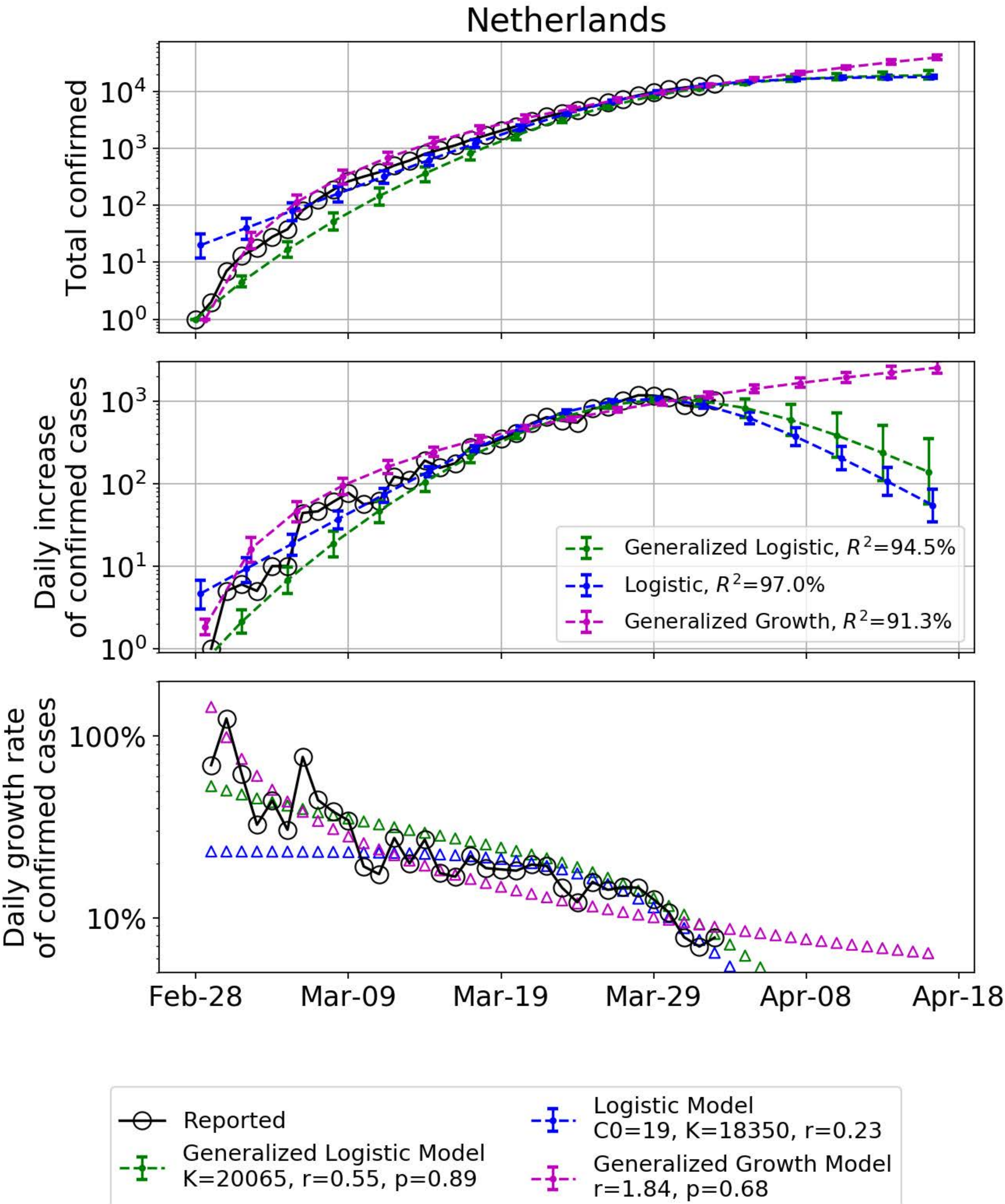
Switzerland
Total confirmed
10^0
10^1
10^2
10^3
10^4
Daily increase of confirmed cases
10^0
10^1
10^2
10^3
Generalized Logistic, R^2=89.3%
Logistic, R^2=88.8%
Generalized Growth, R^2=79.8%
Daily growth rate of confirmed cases
100%
10%
Feb-26
Mar-07
Mar-17
Mar-27
Apr-06
Apr-16
Reported
Generalized Logistic Model
K=24808, r=0.68, p=0.86
Logistic Model
C0=79, K=22750, r=0.20
Generalized Growth Model
r=3.45, p=0.60



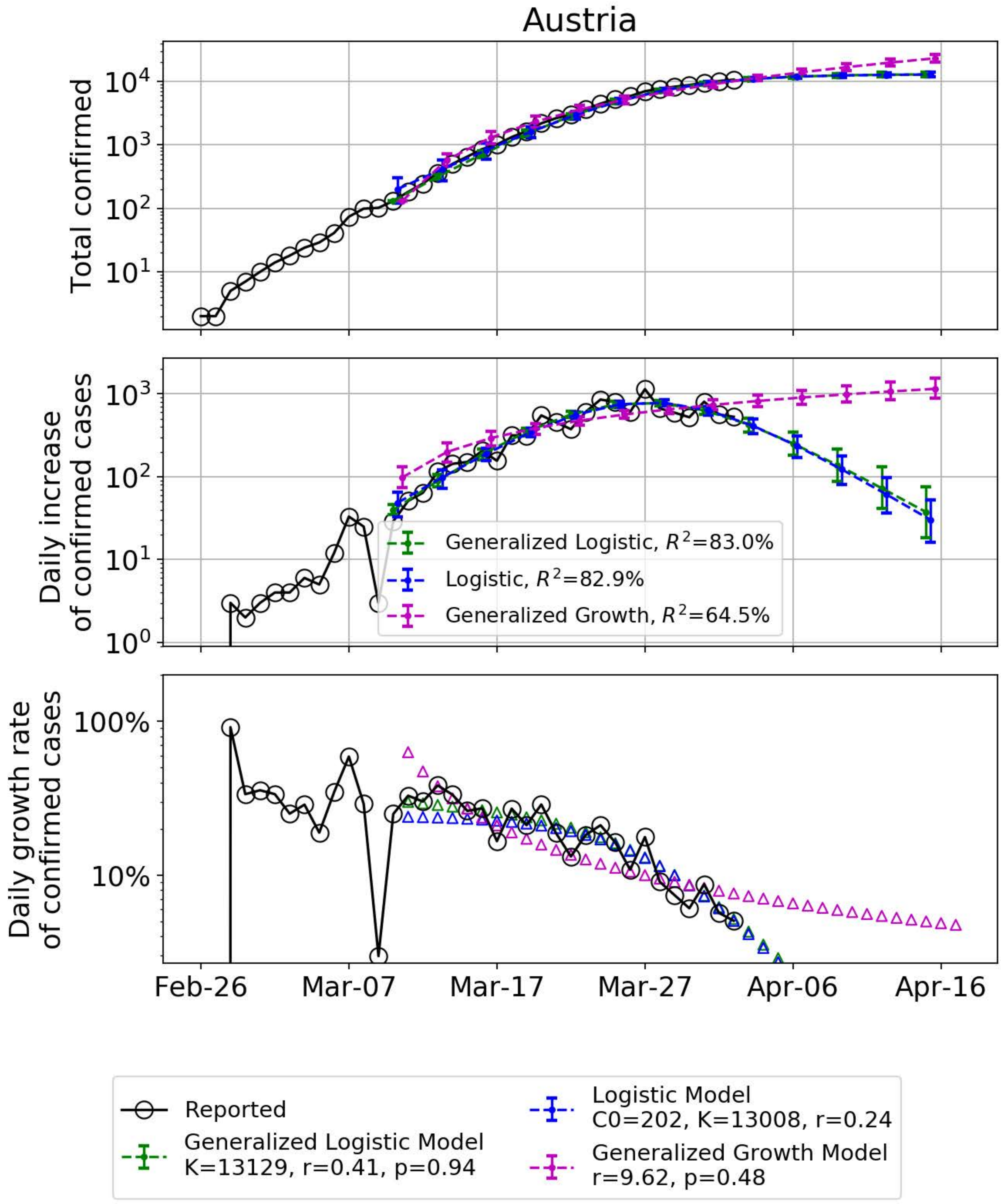
Turkey
Total confirmed
Daily increase of confirmed cases
Daily growth rate of confirmed cases
Generalized Logistic, R^2=94.5%
Logistic, R^2=94.5%
Generalized Growth, R^2=88.8%
Mar-12
Mar-17
Mar-22
Mar-27
Apr-01
Apr-06
Apr-11
Apr-16
Reported
Generalized Logistic Model
K=22665, r=0.67, p=0.94
Logistic Model
C0=47, K=22149, r=0.40
Generalized Growth Model
r=3.13, p=0.70



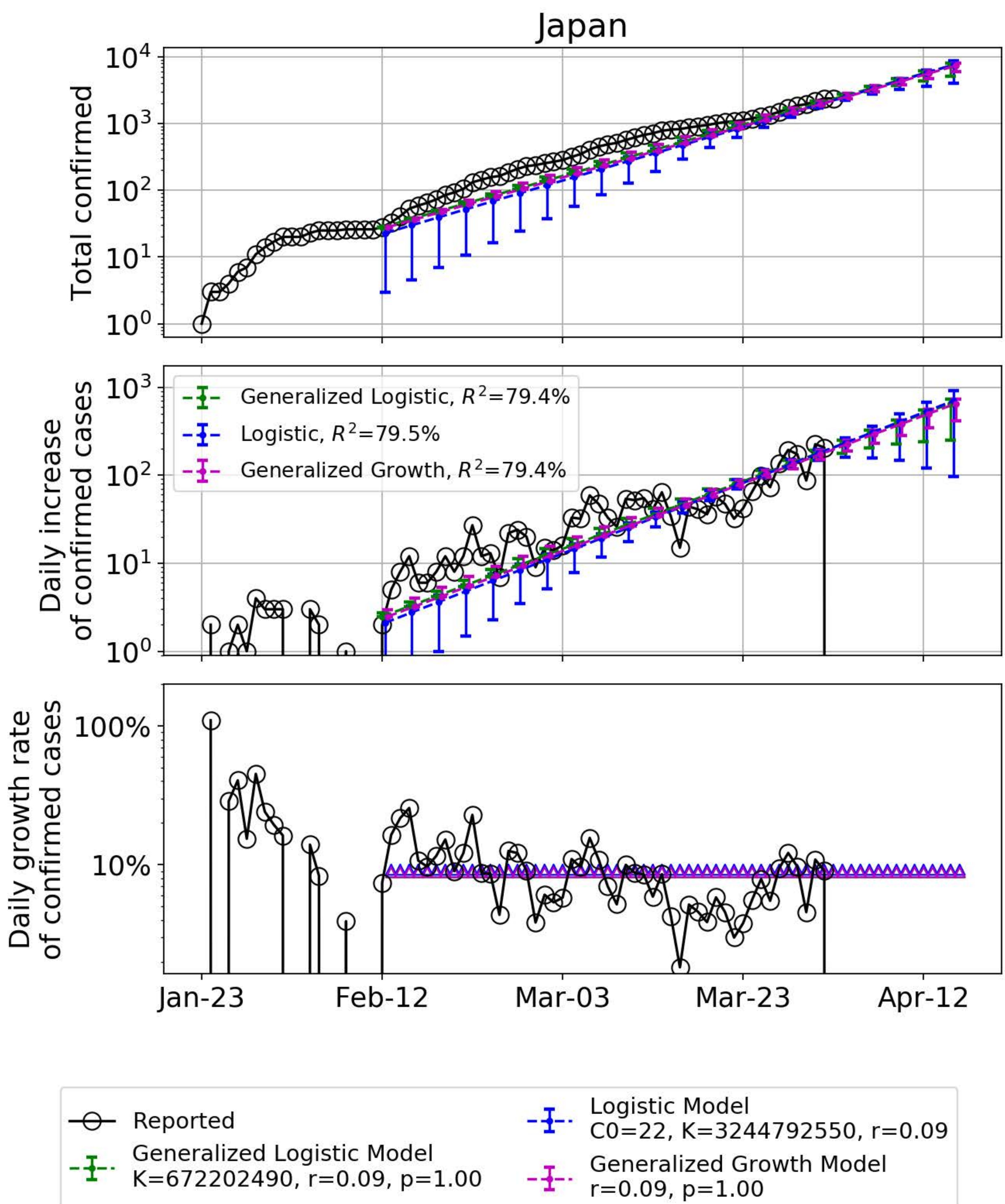
Belgium
Total confirmed
Daily increase of confirmed cases
Daily growth rate of confirmed cases
Generalized Logistic, R^2=85.9%
Logistic, R^2=86.7%
Generalized Growth, R^2=78.1%
Feb-04
Feb-14
Feb-24
Mar-05
Mar-15
Mar-25
Apr-04
Apr-14
Reported
Generalized Logistic Model
K=19637, r=0.27, p=1.00
Logistic Model
C0=4, K=16524, r=0.34
Generalized Growth Model
r=1.46, p=0.73



Netherlands
Total confirmed
Daily increase of confirmed cases
Daily growth rate of confirmed cases
Generalized Logistic, R^2=94.5%
Logistic, R^2=97.0%
Generalized Growth, R^2=91.3%
Feb-28
Mar-09
Mar-19
Mar-29
Apr-08
Apr-18
Reported
Generalized Logistic Model
K=20065, r=0.55, p=0.89
Logistic Model
C0=19, K=18350, r=0.23
Generalized Growth Model
r=1.84, p=0.68



Austria
Total confirmed
Daily increase of confirmed cases
Generalized Logistic, R^2=83.0%
Logistic, R^2=82.9%
Generalized Growth, R^2=64.5%
Daily growth rate of confirmed cases
Feb-26
Mar-07
Mar-17
Mar-27
Apr-06
Apr-16
Reported
Generalized Logistic Model
K=13129, r=0.41, p=0.94
Logistic Model
C0=202, K=13008, r=0.24
Generalized Growth Model
r=9.62, p=0.48



Japan
Total confirmed
Daily increase of confirmed cases
Daily growth rate of confirmed cases
Generalized Logistic, R^2=79.4%
Logistic, R^2=79.5%
Generalized Growth, R^2=79.4%
Jan-23
Feb-12
Mar-03
Mar-23
Apr-12
Reported
Generalized Logistic Model
K=672202490, r=0.09, p=1.00
Logistic Model
C0=22, K=3244792550, r=0.09
Generalized Growth Model
r=0.09, p=1.00