medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Analysis of Covid-19 Data for Eight European Countries and the United Kingdom Using a Simplified SIR Model

Gyan Bhanot[1,2*] , Charles DeLisi[3]

[1] Department of Physics and Astronomy, Rutgers University, Piscataway, NJ, 08854, USA; Department of Molecular Biology and Biochemistry, Rutgers University, Piscataway, NJ, 08854, USA; Rutgers Cancer Institute of New Jersey, New Brunswick, NJ, 08903, USA
[2] Moores Cancer Center, University of California San Diego, La Jolla, CA 92103, USA.
[3] Department of Biomedical Engineering, Boston University, Boston, MA 02215, USA

* Gyan Bhanot, PhD (corresponding author)
136 Frelinghuysen Road, Busch Campus, Rutgers University, Piscataway, NJ 08854, USA
Phone: (848) 391-7508
Fax: (732) 235-5331
Email: gyanbhanot@gmail.com

Charles DeLisi, PhD
24 Cummington Mall
Boston University, Boston, Ma 02215
Phone 617 875 5138
Email: DeLisi@bu.edu

Abstract
Understanding the characteristics of the SARS-Cov-2/Covid-19 pandemic is central to developing control strategies. Here we show how a simple Susceptible-Infective-Recovered (SIR) model applied to data for eight European countries and the United Kingdom (UK) can be used to forecast the descending limb (post-peak) of confirmed cases and deaths as a function of time, and predict the duration of the pandemic once it has peaked, by estimating and fixing parameters using only characteristics of the ascending limb and the magnitude of the first peak. As with all epidemiological analyses, unanticipated behavioral changes will result in deviations between projection and observation. This is abundantly clear for the current pandemic. Nonetheless, accurate short-term projections are possible, and the methodology we present is a useful addition to the epidemiologist's armamentarium. Since our predictions assume that control measures such as lockdown, social distancing, use of masks etc. remain the same post-peak as before peak, deviations from our predictions are a measure of the extent to which loosening of control measures have impacted case-loads and deaths since the first peak and initial decline in daily cases and deaths. The predicted and actual case fatality ratio, or number of deaths per million population from the start of the pandemic to when daily deaths number less than five for the first time, was lowest in Norway (pred: 44 $\pm$ 5 deaths/million; actual: 36 deaths/million) and highest for the United Kingdom (pred: 578 +/- 65 deaths/million; actual 621 deaths/million). The inferred pandemic characteristics separated into two distinct groups: those that are largely invariant across countries, and those that are highly variable. Among the former is the infective period, T_L ($\overline{T_L}$ = 16.3 $\pm$ 2.7 days); the average time between contacts, T_R ($\overline{T_R}$ = 3.8 $\pm$ 0.5) days

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

and the average number of contacts while infective , R ($\overline{R} = 4.4 \pm 0.5$). In contrast, there is a highly variable time lag T_D between the peak in the daily number of confirmed cases and the peak in the daily number of deaths, ranging from a low of $T_D = 2{,}4$ days for Denmark and Italy respectively, to highs of $T_D = 12$, 15 for Germany and Norway respectively. The mortality fraction, or ratio of deaths to confirmed cases, was also highly variable, ranging from low values 3%, 5% and 5% for Norway, Denmark and Germany respectively, to high values of 18%, 20% and 21% for Sweden, France, and the UK respectively. The probability of mortality rather than recovery was a significant correlate of the duration of the pandemic, defined as the time from 12/31/2019 to when the number of daily deaths fell below 5. Finally, we observed a small but detectable effect of average temperature on the probability α of infection per contact, with higher temperatures associated with lower infectivity. Policy implications of our findings are also briefly discussed.

Keywords: Epidemic modelling, SARS-Cov-2/Covid-19 pandemic, SIR model, predictions.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

I. Introduction

Coronaviruses are large, enveloped, single-stranded RNA viruses which are widespread in animals and usually cause only mild respiratory illnesses in humans [1-5]. However, in 2003, a new coronavirus SARS-CoV emerged, which caused a life-threatening respiratory disease, with a fatality rate of almost 10% [6,7]. Unfortunately, after an initial burst of interest in development of treatment options, interest in this virus waned. The emergence of a novel coronavirus SARS-CoV-2, identified in January 2020 as the likely causative agent of a cluster of pneumonia cases which first appeared in Wuhan, China in December 2019, has since caused a worldwide pandemic [8-13]. SARS-CoV-2 is the seventh known coronavirus to cause pathology in humans [1]. The associated respiratory illness, called COVID-19, ranges in severity from a symptomless infection [8], to common-cold like symptoms, to viral pneumonia, organ failure, neurological complications and death [9-11]. While the mortality in SARS-CoV-2 infections is lower than in SARS-CoV [9-12], it has more favorable transmission characteristics, a higher reproduction number [13], and as we will show, a long latency period and asymptomatic infective phase.

The governments of several countries have taken significant measures to slow the infection rate of Covid-19, such as social distancing, quarantine, identification, tracking and isolation. However, there is no uniform policy, some governments have reacted later than others and some (e.g. Sweden) made a deliberate decision to keep the country open, leaving counter-measures up to individuals.

A large amount of consistent public data is now available on the number of tests performed, the number of confirmed infected cases, and the number of deaths in different contexts, such as locations and health conditions [14]. These provide important sources of information for the development and testing of models that can identify pandemic characteristics affecting viral dynamics, and guide public policy by predicting the impact of various interventions [15].

All data, of course, have limitations, and it's ultimately the completeness and quality of data that limit the success of models. It is well known that confirmed infected cases seriously underestimate the actual number of infections [16,17]: not everyone who is infected is symptomatic, and not everyone who dies from the disease has been tested [18]. Even the number of reported deaths may be underestimated because of co-mortalities; i.e. COVID-19 increases susceptibility to other diseases and conditions [19]. Moreover, the virus can be transmitted by asymptomatic individuals – who comprise a substantial portion of the infected population [20] —militating against accurate estimates of transmission probability. Nonetheless as indicated in [21] and by our own verified forecasts, models can provide useful information.

Dynamical (mechanistic) models, such as the one presented here, have been used for forecasting (meaning that once initial conditions are set, there are no changes in the model) and for projecting (the outcome is changed by intervention strategies). For example, projections and forecasting models of various types have been used as early as February, 2020 to determine a reproductive number [13]. More generally multiple research groups have used them to estimate Case Fatality Ratios (CFRs) [22], to forecast and project the need for hospital beds [23] and to project and forecast mortality [24]**.** More specifically, among the many applications to COVID-19, four variable Susceptible-Exposed-Infective-Recovered models have been used to project the

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

impact of social distancing on mortality [25], three variable Susceptible-Infective-Recovered models to estimate case fatality and recovery ratios early in the pandemic [26], and a time delayed SIR has been used to evaluate the effectiveness of suppression strategies [27]. One of the most ambitious dynamical models, which includes 8 state variables and 16 parameters was fruitfully applied to evaluate intervention strategies in Italy, in spite of the fact that parameter identifiability could not be assured [28].

Our simplified version of the original SIR model [29] differentiates itself from the studies done so far by a unique methodological approach to rigorous identification of parameters, and by making a number of useful predictions very easily, such as the duration of the pandemic in different countries and case fatality ratios, using only characteristics of the increasing portion of verified cases vs time, and the amplitude of the peak in the daily cases. Although it is sometimes said that using daily mortality data is more reliable than daily verified cases, we find that at least in the countries we analyzed, the patterns for these two quantities are virtually identical, differing only in amplitude and a rigid time translation.

It is important to note that extrapolating our model predictions for daily cases and deaths past the peak assumes that control measures such as lockdown, social distancing, use of masks etc and care given to Covid-19 patients remained the same after the peak in daily cases as before the peak. Observed deviations from our predictions are a measure of the extent to which loosening of control measures or changes in care of infected individuals impacted case-loads and deaths past the peaks in the number of cases and deaths.

II. Methods:

We model the Covid-19 pandemic using a simplified version of the SIR model [29], which partitions the population into three compartments, Susceptibles (S), Infectious (I) and Removed R: Recovered or Dead after being infected. This and other models to study the global spread of diseases have been used in a variety of contexts (For some recent reviews, see [30-32]).

So far, the Covid-19 pandemic, at least in the developed countries in Europe where we will apply this model, seems to have the following dynamics: After being infected, an individual remains able to infect others for an average of T_L days. After a time T_L, the infected individual becomes sick, gets tested, is identified as infected and is removed from the pool by quarantine or hospitalization. Thus, in our context, the SIR model dynamics can be defined as follows: At t=0, there is a pool of N interacting individuals, almost all of whom are in the S compartment, except for the few infected cases in the I compartment. The R compartment is empty at t=0. Over time, individuals move from S to I and from I to R. In R, they either recover or die. Since the Recovered pool is populated only from the Infected pool, on average, the number removed each day must equal the number infected sometime in the past; i.e. the two are related by a fixed time displacement and a "mortality probability" factor. We assume that the number of deaths and the number of individuals recovered each day are proportional to the number Removed each day by fixed probabilities, that remain invariant over the course of the epidemic; i.e. that the number dead or recovered each day are proportional to the number infected on some previous day, with different time delays and probabilities.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The Model:

Let,

$X_1\ (t)\ =$ number of Susceptible individuals at time t, (1)
$X_2\ (t)\ =$ number of Infected individuals at time t, (2)
$X_3\ (t)\ =$ number of individuals that are Removed at time t (3)
$X_1\ (t) +\ X_2\ (t) +\ X_3\ (t) =\ \ N$ (4)

A fraction δ of the infected individuals will die after being identified as infected. On average, there will be a time delay T_D between when a person is identified to be infected (tests positive) and when he/she dies of the disease. T_D will depend on a variety of factors, such as quality of care, age, severity of disease, co-morbidities, immune status etc.

Under these assumptions, the number of deaths $X_4(t)$ at time t is related to the number of Infected cases X_2 by:

$X_4(t) =\ \delta\, X_2(t\ - T_D) =$ number of individuals that died on day t (5a)

Similarly, the number of Recovered at time t will be:

$X_5(t) =\ \beta\, X_2(t\ - T_{Re}) =$ number of individuals that recovered on day t (5b)

Let,

α be the transmission rate, the probability of infection per day per contact. (6)
and let γ be the rate at which individuals leave the infected population = $\frac{1}{T_L}$, (7)
with T_L being interpreted as the latency, or the average time interval during which an infected individual can infect a naive individual.

The equations governing the dynamics are:

$$\frac{dX_1}{dt} =\ -\alpha X_1 X_2 \quad (8)$$

$$\frac{dX_2}{dt} =\ \alpha X_1 X_2 -\ \gamma X_2 \quad (9)$$

The initial conditions at $t = 0$ are:

$$X_1(0) = (N - a) \sim N \text{ and } X_2(0) = a; \text{with } a \ll N \quad (10)$$

An equation relating the state variables X_1 and X_2 can be obtained by dividing (8) by (9) and integrating. This gives,

$$X_2(t) =\ N + \frac{N}{R}\log\left(\frac{X_1(t)}{N}\right) -\ X_1(t) \quad (11)$$

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

where $R \equiv \frac{\alpha N}{\gamma}$ (12)

Hence,

$$\frac{dX_1}{dt} = -\alpha X_1 \left[N + \frac{N}{R} \log\left(\frac{X_1(t)}{N}\right) - X_1(t) \right] \quad (13)$$

From (9) the maximum in X_2 is at $X_1 = \frac{\gamma}{\alpha} = \frac{N}{R}$ (14)

Substituting this into (13) gives:

Maximum value of $X_2 \equiv P = N - \frac{N}{R}[1 + \log(R)]$ (15)

At $t = \infty, X_2 = 0$. Hence, from (13), we get:

$$R = -\frac{\log(S_1(\infty))}{[1 - S_1(\infty)]}, \text{ where } S_1(\infty) = \frac{X_1(\infty)}{N} \quad (16)$$

as the fraction of susceptible individuals at t = ∞.

When $S_1(\infty) \to 1, \log(S_1(\infty)) \sim -(1 - S_1(\infty))$ and $R \to 1$

This says that when nobody is infected, $S_1(\infty) = 1$, i.e. there is no pandemic. This corresponds to $R = 1$. It is easy to show that if $R < 1$, there are no solutions to (16) that satisfy $0 \leq S_1(\infty) \leq 1$.

Fitting the Model to Data

We focus on four North European countries, Netherlands, Denmark, Sweden and Norway, denoted by EN, four South European countries, France, Italy, Spain and Germany, denoted by ES, and the United Kingdom (UK). Data for the number of confirmed cases and deaths were obtained from https://ourworldindata.org/coronavirus-source-data, from an EU agency established in 2005 and based in Stockholm with the aim to strengthen Europe's defense against infectious diseases. The data for the number of tests was obtained from https://ourworldindata.org/coronavirus-testing [33].

We note that the data identify the daily number of Confirmed cases, whereas the SIR model we described above requires the number of Infected cases. However, the number of confirmed cases at time t are derived from the infected cases at some previous time t', where the time lag between t and t' is fixed but unknown. Since Equations (1-5) are invariant under a time translation, we can use the data for daily confirmed cases to represent X_2 with the understanding that there is an implicit shift in time between a person becoming infected and being identified as such in the data. This shift in time would depend on how quickly infected individuals are identified and included in the case count, and hence may vary from country to country. However, we assume that within a

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

given country, this time interval is fixed on average. With this caveat, in the results below, we will use the confirmed daily cases in our analysis to represent $X_2(t)$, and use either the term "infected" or "confirmed" when discussing X_2.

The first date for which data was available for these countries was 12/31/2019, which we denote as day number 0 in our analysis and in the plots to follow. Among the countries considered here, the earliest cases were identified in France on 1/25/2020, which corresponds to day 25. In the other countries, the earliest cases were identified on the following days, counting from 12/31/2019: Netherlands: day 59, Denmark: day 58, Sweden: day 32, Norway: day 58, UK: day 31, Spain: day 32, Germany: day 28, Italy: day 31.

The available data includes
(a) The cumulative number $I(t)$ and daily number $X_2(t)$ of confirmed cases (17)
and
(b) The cumulative number $D(t)$ and daily number $X_4(t)$ of deaths (18)

These are related by:

$$X_2(t) = I(t) - I(t-1) \tag{19}$$

and,

$$X_4(t) = D(t) - D(t-1)\,. \tag{20}$$

It is worth noting that the dynamics of the number removed each day and the total number of removed do not enter our analysis explicitly. Some connection to the full SIR model [29] can be made by noting that the total number of removed individuals increases asymptotically at large times to $N(1 - S_1(\infty))$.

We determine and analyze the following parameters country by country: $\boldsymbol{N}$, as defined in (4); $\boldsymbol{\alpha}$, the transmission rate; i.e. the number of infections per day per contact; $\boldsymbol{\gamma}$, the average rate at which individuals leave the infected pool; $\boldsymbol{R}$, the average number of transmissions per individual; $\boldsymbol{T_R}$, the time between transmissions; $\boldsymbol{\delta}$, the fraction of individuals in the infected pool who will, on average, die $\boldsymbol{T_D}$ days later; $\boldsymbol{\beta = 1 - \delta}$, the fraction of individuals in the infected pool who will, on average, recover $\boldsymbol{T_{Re}}$ days later; and $\boldsymbol{T_L}$, the infective period. We also determine the duration of the infection, defined as the number of days from 12/31/19 until the number of deaths drops below 5 per day.

The parameters N, α, γ, R were obtained using (5), (8) and (9) to do numerical fits. Because of the definition of R (see (12)), only three of the parameters, N, α, γ, R, need to be determined from the data. Consequently, we proceed as follows:
1. Using (12), we define α in terms of N, γ, R. This eliminates α.
2. Estimating P = maximum value of $X_2(t)$ from the data, we determine N in terms of R using (15). This eliminates N.
3. $\gamma(R-1)$ is determined as the coefficient of t in the exponential rise of $X_2(t)$ for small t (see Appendix A). This eliminates γ.
4. Using a numerical solver, we vary R to fit the observed data for $X_2(t)$.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

We emphasize that in determining the parameters, we are only using the ascending limb and the peak in the data for $X_2(t)$ and use these to predict how $X_2(t)$ will evolve in time past the peak. Using the data for $X_4(t)$ and the fitted model, we can also determine T_D and δ using (5a) by a simple translation and scaling of the fitted model for $X_2(t)$.

Furthermore, once N, α, γ, R are determined, (16) determines $S_1(\infty) = \frac{X_1(\infty)}{N}$, the fraction of the pool of interacting individuals who are naive (uninfected) at the end of the pandemic.

Some other useful parameters we obtain are:

The infective period: $T_L = \frac{1}{\gamma}$, (21)

the average time between transmissions: $T_R = \frac{1}{N\alpha}$, (22)

and the average number of transmissions while infective = $N_C = \frac{T_L}{T_R} = \frac{N\alpha}{\gamma} = R$ (23)

Data and Fitting Methodology

Although data on deaths is unambiguous, the data for the number of cases is trustworthy only when a sufficient number of tests are performed. It is therefore important to determine whether adequate testing was done to ensure the reliability of case data. Figure 1a and 1b show the cumulative number of tests performed in the countries analyzed, starting from 12/31/2019. We see that in the EN countries, the ratio of cumulative tests to cumulative cases always exceeded five, whereas in the ES countries and the UK, it always exceeded three. Consequently, we expect that the reported number of cases is reliable. It is also important to note that while replicate testing provides some assurance, it leaves unaddressed the problem of quality (reliability), breadth (the percentage of the population reached) and speed (the elapsed time between infection and identification of disease) of testing.

One interesting observation from the testing data (Figure 1a, 1b) was that initially, *public concern spread more rapidly than the virus*. As the disease spread and testing ramped up, the initial value of T_C (number of tests/case) was quite large (over 650 in the UK), suggesting that the exponential growth in infections was slower than the exponential public awareness of the disease, resulting in the testing of a large number of people with no disease. Later, as appropriate protocols to identify diseased individuals were established, T_C decreased then flattened out when the testing capacity adequately accounted for new cases, and finally rose again as the pandemic waned.

The values of $X_2(t)$ and $X_4(t)$ were extracted from the data for the cumulative number of cases and the cumulative number of deaths. To reduce fluctuations in these quantities, the data was averaged over 7 days. Once B = $\gamma(R-1)$ and P were determined from the exponentially increasing region and the peak in $X_2(t)$ respectively (see discussion in previous section), the only undetermined parameter was R. Several values of R in the range $R = 1.5 - 6.0$ were then tested using the following procedure: For each parameter set, we fitting these data for $X_2(t)$ from small t to the peak and beyond, by finding numerical solutions of (8) and (9) using the Matlab Solver *myode2* to determine $[X_1(t), X_2(t)]$ as a function of time, with the initial conditions, $[X_1(t_0) = N - a, X_2(t_0) = a]$, starting from a value t_0 of t such that $X_2(t_0) = a \sim 10$.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

To determine an error on the fitted parameters, the parameters were varied until a range of values was found that fit the data for $X_2(t)$ including fluctuations. Finally, δ, T_D were emperically determined (see eq. (5a)) by scaling and shifting the numerical fits for $X_2(t)$ to fit the data for $X_4(t)$. Using the fitted parameters, the numerical solutions for $[X_1(t), X_2(t)]$ were extended to estimate the number of days to the end of the pandemic, *which we define as when the number of daily deaths are less than 5*. The fitted solutions for $[X_2(t), X_4(t)]$ were also used to compute the predicted total cases per million and the predicted total deaths per million population and compared to their actual values in each country. Finally, the values of T_L and T_R were determined from the fitted data using (21) and (22) respectively.

III. Results and Discussion

It is important to note that the parameter values we determine are derived solely from: (i) the ascending (small t) limb of $X_2(t)$ and (ii) the first peak value of $X_2(t)$. Using the fitted parameters to derive results past the peak allows us to: (i) predict how the pandemic would have continued beyond the peak, if control measures such as social distancing, use of PPEs and masks etc. had remained the same as before the peak; (ii) determine deviations of the actual data from our model predictions to estimate the effect of changes in control measures post peak to pandemic dynamics.

The results obtained for $N, \alpha, \gamma, R, \delta, T_L, T_R, T_D$ are given in Tables A. Figures 2 and 3 and Supplementary Figure S1a-i shows the data and fits of our model for X_2 and X_4 respectively. We note that for Sweden and the UK, there is no clear first peak in $X_2(t)$ but rather a plateau and a subsequent increase in the number of daily cases. For these countries, we estimated the value of P (Eq. 15) from the plateau in Fig. 2c and Fig 2e respectively. For all the other countries, there was a clear first peak in $X_2(t)$ whose amplitude was used to estimate P. Unfortunately, for all countries, even those that had a well defined first peak in followed by a decline over time, $X_2(t)$ showed significant increase soon after the peak. In some cases, subsequent increases in daily deaths surpassed these first peaks by factors of 2-6 (Figure 1, S1). These suggest that in spite of significant early efforts to control infection rates using social distancing, quarantine, use of masks, testing and other containment measures, there was poor success in limiting the spread of the disease past the first peak in X_2 (Figure 2 a-i), most likely because of a combination of premature lifting of quarantine, inadequate use of facemasks and difficulties in limiting physical contacts, exacerbated by the advent of summer weather and holiday travel.

In contrast, for the number of daily deaths X_4 (Figure 3 a-i), there were significantly smaller increases past the peak for all countries, suggesting that measures to save lives worked better than measures to contain the number of infections. This is likely due to the efficiency of the health care systems in these countries and the competence and professionalism of health care workers.

Results for T_D, δ, T_L, the time delay between case identification and death and the time interval T_R between transmissions while infective:

The fraction δ of identified cases who died after a time interval T_D (Figure 4a, Table A) also shows significant variation by country. Norway, Germany and Denmark have the smallest values: $\delta =$

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

$0.03, 0.050$ and 0.050 respectively, and the UK, France and Sweden have the highest: $\delta = 0.21, 0.20$ and 0.18 respectively. Assuming that most of the deaths occurred in hospitals, the average time T_D from case identification to death for those who died was highest (12, 15 days) for Germany and Norway respectively and lowest (2, 4 days) for Denmark and Italy respectively (Figure 4b, Table A).

It is significant that in each country, the relation between the daily number of cases and the daily deaths is described by just two parameters, δ and T_D (Eq. 5a), for the entire duration of the pandemic. This is not required a priori. For example, pressure on resources during a peak period of infectivity might have caused a transient increase in the number of deaths per day relative to the number of recovered per day. In such a situation, a single probability δ and a single value of T_D need not have sufficed for the entire epidemic. The fact that there is little evidence of variations in these parameters over the pandemic, to within the quality of the data, suggests that although the effectiveness of life saving measures may differ between countries, these measures seem to be relatively insensitive to changes in the case burden.

The average infective period $\overline{T_L} = 1/\gamma$ was approximately constant for all countries, with an average value: $\overline{T_L} = 16.3 \ +/- \ 2.7$ days. The average time between transmissions $\overline{T_R} = \frac{1}{N\alpha}$ was also remarkably uniform across countries, averaging : $\overline{T_R} = 3.8$ +/- 0.5 days. Finally, the average number of transmissions while infective or $\overline{N_C} = \frac{\overline{T_L}}{\overline{T_R}}$, which is also the average value of R, varied only in a narrow range for all countries, averaging: $\overline{N_C} = \bar{R} = 4.4$+/- 0.5.

The duration of the pandemic, the case fatality ratio (deaths/million) and cases/million.

We define the end of the pandemic as the day when daily deaths becomes less than 5 for the first time after first peaking. The duration of the pandemic in each country is shown in Figure 4c as the number of days from 12/31/2019 till the day that daily deaths became < 5. With this definition, the predicted and actual dates are (Table A) : Netherlands: (6/18 +/- 10 days, 6/17), Denmark: (5/11 +/- 4 days, 5/10);, Sweden: (6/28 +/- 10 days, 7/26), Norway: (4/26 +/- 3 days, 4/28), UK: (8/5 +/-13 days, 8/21), Spain: (6/29 +/-10 days, 6/29), Germany: (6/24 +/- 9 days, 7/14), France: (7/12 +/-12 days, 7/19), Italy: (7/28 +/- 15 days 8/2). Note that the model predictions for pandemic duration are in quite good agreement with observation.

The measured case fatality ratio or the total number of deaths per million population (Table A) ranged from low values of 36 and 89 for Norway and Denmark, to high values of 621, 604 and 582, 562 for the UK, Spain, Italy and Sweden respectively. The model predicted values for this quantity are in good agreement with the actual values (Figure 4d).

The measured total number of cases per million population (Table A) ranged from low values of 1367 and 1774 for Norway and Denmark to high values of 7375, 5293, 4806 and 4092 for Sweden, Spain, Sweden, the UK and Italy respectively. In UK and Spain, the model predictions for this quantity were in sharp disagreement with the actual numbers (Figure 4e), likely the result of a significant relaxation of containment measures after the peak in the daily cases was reached.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

SARS-Cov-2 may transmit less effectively at higher temperatures.

As was also noticed in [34], an interesting finding was a weak "Temperature Effect" on the value of the infectivity parameter α. Figure 4f shows the average temperature in February 2020 for the principal cities in each country versus the fitted value of α, and suggests that SARS-Cov-2 may transmit less efficiently at higher temperatures. However, this result is far from final and needs further testing and validation because of many confounding factors, such as humidity, use of air conditioning and exhaust fans, crowding, population density etc. that may also have significant effects on α.

In summary, the simple extension of the original SIR model [29] that we propose in this paper is able to identify the full set of pandemic parameters and fully characterize all three compartments S, I and R. It would be useful to try to understand what policy measures would result in improved values for some of these parameters. However, it is also important to recognize that some of the parameters cannot be controlled by policy changes, and some that can be will not all respond to the same strategy. For example, T_L can be shortened by increasing the frequency and number of tests, since both these strategies increase the number of people removed from the infecting pool. On the other hand, the probability that an infected individual will die as opposed to recover might be affected by resources such as ventilators and masks. These considerations lead us to ask whether the degree of correlation between the parameters of the model, and features of the pandemic such as duration and deaths per million population, can account for the wide intercountry variation in pandemic features. A low degree of correlation for a particular parameter would indicate that it had little influence on variation of the feature, whereas a high degree of correlation would be informative.

To estimate some dependencies that may have useful policy implications, we computed the Spearman Correlation between the derived parameters T_L, Nα, and δ and some metrics of value in determining policy – namely " PD = pandemic duration" and "CFR = case fatality ratio or deaths per million." We found that $N\alpha$ was not correlated with these metrics (Correlation with PD: -0.20, p-value 0.60; Correlation with CFR: -0.02, p-value 0.98). However, T_L was somewhat correlated (Correlation with PD: 0.64, p-value 0.07; Correlation with CFR: 0.533, p-value 0.15) and δ was highly correlated (Correlation with PD: 0.86, p-value 0.005; Correlation with CFR: 0.76, p-value 0.02). These results suggest that the optimum method to decrease both the duration of the pandemic and the case fatality ratio would be to decrease the infective period T_L and δ, which are achievable by early case identification, contact tracing and quarantine (which would reduce T_L) and improving quality of care for identified cases (which would reduce δ). Whereas these relationships may seem obvious, and indeed have been used as mechanisms to limit the impact of the pandemic in many countries, the fact that they are also readily identifiable in the data is, in our opinion, quite interesting.

The value of $N\alpha$, the average number of new infections per day, has clear implications for the time it will take a randomized trial to establish therapeutic efficacy. While a statistical analysis is outside the scope of this manuscript, we note that the wide inter-country dispersion in this parameter indicates that the choice of population could have a pronounced effect on the time to establish efficacy, and it might therefore be best to select a test population deliberately from a country or

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

countries with large and similar values of this parameter, rather than mix populations from different countries with no consideration of the average rate of infectivity.

Declarations

Funding and Acknowledgments
GB was partly supported by grants from M2GEN/ORIEN, DoD/ KRCP (KC180159) and NIH/NCI (1R01CA243547-01A1). He thanks Professors Pablo Tamayo and Jill Mesirov for their kind hospitality at UC San Diego during his sabbatical year 2019-2020 when this work was done. GB thanks Kevin Raines, Sebastian Doniach, Alain Billoire and Carl Dietrich Foerster for many helpful discussions and suggestions.

Conflicts of interest/Competing interests
The authors declare no conflict of interests.

Ethics approval
Not applicable

Consent to participate
Not applicable

Consent for publication
Not applicable

Availability of data and material
The data used in this paper were all derived from public sources. Links to these data are included in the paper.

Code availability
The Matlab codes used to analyze the data along with all data files will be provided on request. Please send an email to gyanbhanot@gmail.com if you want these codes.

Author Contributions
GB: Idea development, analysis, manuscript.
CD: Idea development, analysis, manuscript.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure Legends

Figure 1: Ratio of the cumulative number of tests to number of identified Covid-19 positive cases as a function of time: (a) In the EN countries and **(b)** in the UK and ES countries. The dashed lines in **(a)** and **(b)** represent 5 and 3 tests/case respectively.

Figure 2: Parameter fits for the number of daily cases: Blue circles are observed data for $X_2(t)$, the number of cases per day. Solid lines are fits obtained by solving (8) and (9) using the ODE solver *ode45* in Matlab. The values of the parameters are shown in the insets and represent the fit shown as the solid red line. The method used for the fits was to find $\gamma(R-1)$ from the exponential rise in X_2 at early times (Appendix A), estimate the peak value P of X_2 (which gives the value of N using (15)) and find an R value that best fits the data (red solid line). The black lines represent solver results for R varying by 0.5 from the best fit value.

Figure 3: Number of deaths per day and fits for δ and T_D: Red circles are observed data for $X_4(t)$, the number of deaths per day. The solid lines were obtained by shifting the fits for X_2(t) in Figure 2 forward in time by an amount T_D and scaling the results by δ (see eq. 5a).

Figure 4: Predictions from fits. (a) Best fits for T_D (see eq. 5a and Figure 3), the average number of days from case identification to death, if death occurs. Values range from a high of 15 days for Norway and a low of 4 days for Italy and Denmark. **(b)** Best fits for δ, (see eq. 5a and Figure 3), the fraction of identified cases that died an average of T_D days after case was recorded. Values range from low values 0.032, 0.05 and 0.05 for Norway, Denmark and Germany to high values 0.21, 0.20 and 0.20 for UK, Sweden and France. **(c)** Predicted and actual number of days from 12/31/2019 to when number of daily deaths < 5. Note that for all countries, our predictions are very close to the actual results. This is also evident from the extrapolation of the fits past the peak in Figure 3. This suggests that measures to limit fatalities were just as effective after the peak as before. **(d)** Predicted and actual case fatality ratios (number of deaths per million) from start of the pandemic to when daily deaths < 5. Actual values for the case fatality ratios range from lows of 18 and 76 deaths/million population for Norway and Denmark and highs of 617, 604 and 579 deaths/million population for UK, Spain and Italy. Note that the predicted values from the fits are in good agreement with the actual values for all countries, suggesting consistent medical care throughout the pandemic. **(e)** Predicted and actual cases per million population from start of the pandemic to when daily deaths < 5. Note that now, the predicted numbers are in reasonable agreement for all countries except UK and Sweden, where preventive measures such as partial lockdowns, use of masks and social distancing were not mandated/followed. **(e)** Identification of a possible temperature effect on the infectivity parameter α. The x-axis shows the temperature for February 2020 for the principal cities of the countries studied on a linear scale. The results suggest that the infectivity of SARS-CoV-2 decreases with increasing temperature (see also [34]).

Supplementary Figure Legend

Supplementary Figure S1: Parameter fits for the number of daily cases plotted on a linear scale: These plots are the same as those in Figure 2 except that the y axis is shown in a linear scale, rather than a log scale. This is to show the accuracy of the fits which are shown as solid lines. Blue circles are observed data for $X_2(t)$, the number of cases per day. Solid lines are fits obtained by

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

solving (8) and (9) using the ODE solver *ode45* in Matlab. The values of the parameters are shown in the insets and represent the fit shown as the solid red line. The method used for the fits was to find $\gamma(R-1)$ from the exponential rise in X_2 at early times (Appendix A), estimate the peak value P of X_2 (which gives the value of N using (15)) and find an R value that best fits the data (red solid line). The black lines represent solver results for R varying by 0.5 from the best fit value.

Table A Legend:
Table showing a summary of all results for all 9 countries.

References

1. Thiel V, *Coronaviruses: Molecular and Cellular Biology*, Caister Academic Press 2007, ISBN:978-1-904455-16-5.

2. Rabadan R, *Understanding Coronavirus*, University Printing House, Cambridge, UK (2020), ISBN: 978-1-108-920254.

3. Su S, Wong G, Shi W, Liu J, Lai ACK, Zhou J, Liu W, Bi Y, Gao GF, *Epidemiology, Genetic Recombination, and Pathogenesis of Coronaviruses*, Trends Microbiol. 2016 Jun;24(6):490-502. doi: 10.1016/j.tim.2016.03.003. Epub 2016 Mar 21.

4. Lai MM, Cavanagh D, *The molecular biology of coronaviruses*, Adv Virus Res1997;48:1-100.

5. Masters PS, *The molecular biology of coronaviruses*, Adv Virus Res. 2006;66:193-292.

6. Cherry JD, Krogstad P, *SARS: the first pandemic of the 21st century*, Pediatric Research 2004 Jul;56(1):1-5. Epub 2004 May 19.

7. Peiris J, Guan Y and Yuen K, *Severe acute respiratory syndrome*. Nat Med 10, S88–S97 (2004). https://doi.org/10.1038/nm1143.

8. J. A. Al-Tawfiq, *Asymptomatic coronavirus infection: Mers-cov and sars-cov-2 (covid-19)*, Travel Med Infect Dis, vol. 101608, 2020.

9. Tay MZ, Poh CM, Rénia L, MacAry PA, Ng LFP, *The trinity of COVID-19: immunity, inflammation and intervention*, [published online ahead of print, 2020 Apr 28]. Nat Rev Immunol. 2020;1-12. doi:10.1038/s41577-020-0311-8

10. Xu Z, Shi L, Wang Y, Zhang J, Huang L, Zhang C, Liu S, Zhao P, Liu H, Zhu L, *et. al*., *Pathological findings of covid-19 associated with acute respiratory distress syndrome*, The Lancet Respiratory Medicine, vol. 8, no. 4, pp. 420–422, 2020.

11. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, Xiang J, Wang Y, Song B, Gu X, Guan L, Wei Y, Li H, Wu X, Xu J, Tu S, Zhang Y, Chen H, Cao B, *Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study*. Lancet. 2020 Mar 28;395(10229):1054-1062. doi: 10.1016/S0140-6736(20)30566-3. Epub 2020 Mar 11.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

12. Sørensen MD, Sørensen B, Gonzalez-Dosal R, *et. al., Severe acute respiratory syndrome (SARS): development of diagnostics and antivirals*. Ann N Y Acad Sci. 2006;1067(1):500-505. doi:10.1196/annals.1354.072

13. Inglesby TV, *Public Health Measures and the Reproduction Number of SARS-CoV-2* [published online ahead of print, 2020 May 1], JAMA. 2020;10.1001/jama.2020.7878. doi:10.1001/jama.2020.7878; Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J, *The reproductive number of COVID-19 is higher compared to SARS coronavirus*, Journal of Travel Medicine, Volume 27, Issue 2, March 2020, taaa021, https://doi.org/10.1093/jtm/taaa021.

14. For example, see: https://Ourworldindata.org/coronavirus from The *European Center for Disease Control and Prevention*.

15. Flaxman, S, Mishra, S, Gandy, A et al. *Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe*. Nature 584, 257–261 (2020). https://doi.org/10.1038/s41586-020-2405-7; Davies NG, Kucharski AJ,Eggo RM, Gimma A, Edmunds WJ, *Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study*, The Lancet, July 01, 2020, Volume 5, Issue 7, e375-e385; Ferguson et al., Imperial College Covid-19 Response Team, Report 9(https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-9-impact-of-npis-on-covid-19/).

16. Lu FS, Nguyen AT, Link NB, Davis JT, Chinazzi M, Xiong X , Vespignani A, Lipsitch M, Santillana M, *Estimating the Cumulative Incidence of COVID-19 in the United States Using Four Complementary Approaches*, medRxiv 2020.04.18.20070821; doi: https://doi.org/10.1101/2020.04.18.20070821.

17. COVID-19 Antibody Seroprevalence in Santa Clara County, California
Bendavid E, Mulaney B, Sood N, Shah S, Ling E, Bromley-Dulfano R, Lai C, Weissberg Z, Saavedra-Walker R, Tedrow J, Tversky D, Bogan A, Kupiec T, Eichner, Gupta R, Ioannidis J, Bhattacharya J, *COVID-19 Antibody Seroprevalence in Santa Clara County, California*, medRxiv 2020.04.14.20062463; doi:https://doi.org/10.1101/2020.04.14.20062463

18. https://www.washingtonpost.com/investigations/coronavirus-death-toll-americans-are-almost-certainly-dying-of-covid-19-but-being-left-out-of-the-official-count/2020/04/05/71d67982-747e-11ea-87da-77a8136c1a6d_story.html.

19. https://www.nytimes.com/interactive/2020/06/01/us/coronavirus-deaths-new-york-new-jersey.html

20. Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, Shaman J, *Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2)*, Science, 01 May 2020, Vol. 368, Issue 6490, pp. 489-493.

21. https://covid-19.bsvgateway.org

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

22. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et. al. *Estimates of the severity of coronavirus disease 2019: a model-based analysis*, The Lancet, June 01, 2020, Volume 20, Issue 6, p669-677.

23. https://penn-chime.phl.io/

24. https://www.kff.org/policy-watch/covid-19-models/

25. Branas CC, Rundle A, Pei S, Yang W, Carr BG, Sims S, Zebrowski A, Doorley R, Schluger N, Quinn JW, Shaman J, *Flattening the curve before it flattens us: hospital critical care capacity limits and mortality from novel coronavirus (SARS-CoV2) cases in US counties*, medRxiv 2020.04.01.20049759; doi:https://doi.org/10.1101/2020.04.01.20049759

26. Anastassopoulou C, Russo L, Tsakris A, Siettos C. *Data-based analysis, modelling and forecasting of the COVID-19 outbreak.* PLoS One. 2020, 15(3):e0230405. doi:10.1371/journal.pone.0230405.

27. Casella F. *Can the COVID-19 epidemic be controlled on the basis of daily test reports?* Preprint 2020: https://arxiv.org/abs/2003.06967.

28. Giordano, G., Blanchini, F., Bruno, R. et al. *Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy.* Nat Med 26, 855–860 (2020). https://doi.org/10.1038/s41591-020-0883-7

29. Kermack W, McKendric A (1991), *Contributions to the mathematical theory of epidemics – I. Bulletin of Mathematical Biology*. 53 (1–2): 33–55; ibid. *Contributions to the mathematical theory of epidemics – II. The problem of endemicity*", Bulletin of Mathematical Biology. 53 (1–2): 57–87. doi:10.1007/BF02464424. PMID 2059742; ibid. *Contributions to the mathematical theory of epidemics – III. Further studies of the problem of endemicity*. Bulletin of Mathematical Biology. 53 (1–2): 89–118. doi:10.1007/BF02464425. PMID 2059743.

30. Huppert A, Katriel G. *Mathematical modelling and prediction in infectious disease epidemiology*. Clin Microbiol Infect. 2013;19(11):999-1005. doi:10.1111/1469-0691.12308

31. Brauer F, *Mathematical epidemiology: Past, present, and future*. Infect Dis Model (2017) 2(2):113-127. Published 2017 Feb 4. doi:10.1016/j.idm.2017.02.001

32. Colizza V, Barrat A, Barthelemy M, Valleron AJ, Vespignani A, *Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions*. PLoS Med. 2007;4(1):e13. doi:10.1371/journal.pmed.0040013.

33. Hasell J, Mathieu E, Beltekian D et al. A cross-country database of COVID-19 testing. Sci Data 7, 345 (2020).

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

34. Raines K, Doniach S, Bhanot G, *The transmission of SARS-CoV-2 is likely comodulated by temperature and by relative humidity*, medRxiv 2020.05.23.20111278; doi: https://doi.org/10.1101/2020.05.23.20111278.

All Figures

Fig. 1a

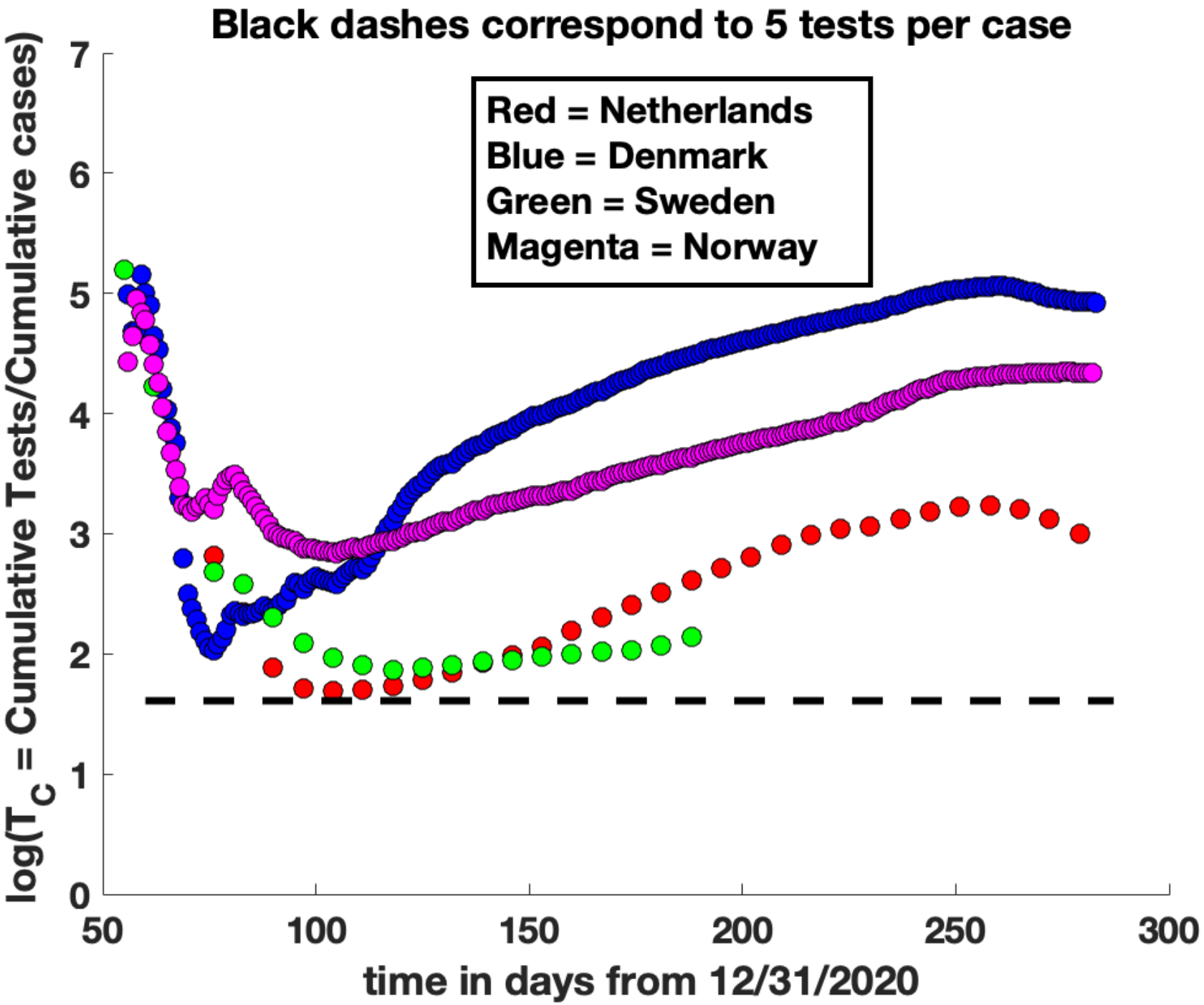

Black dashes correspond to 5 tests per case
Red = Netherlands
Blue = Denmark
Green = Sweden
Magenta = Norway
log(T_C = Cumulative Tests/Cumulative cases)
0
1
2
3
4
5
6
7
50
100
150
200
250
300
time in days from 12/31/2020

Fig. 1b

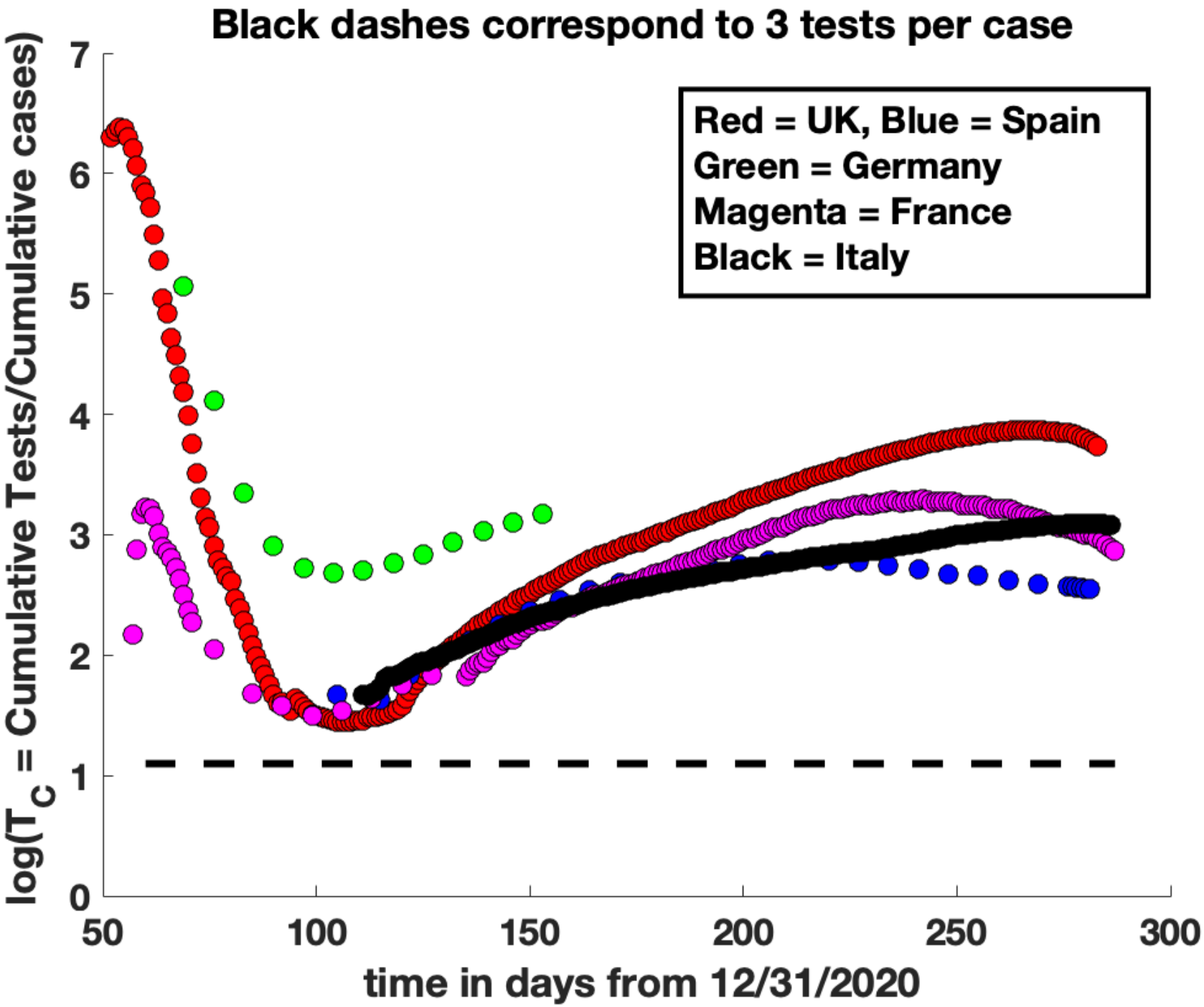

Black dashes correspond to 3 tests per case
Red = UK, Blue = Spain
Green = Germany
Magenta = France
Black = Italy
log(T_C = Cumulative Tests/Cumulative cases)
time in days from 12/31/2020
0
1
2
3
4
5
6
7
50
100
150
200
250
300

Fig. 2a

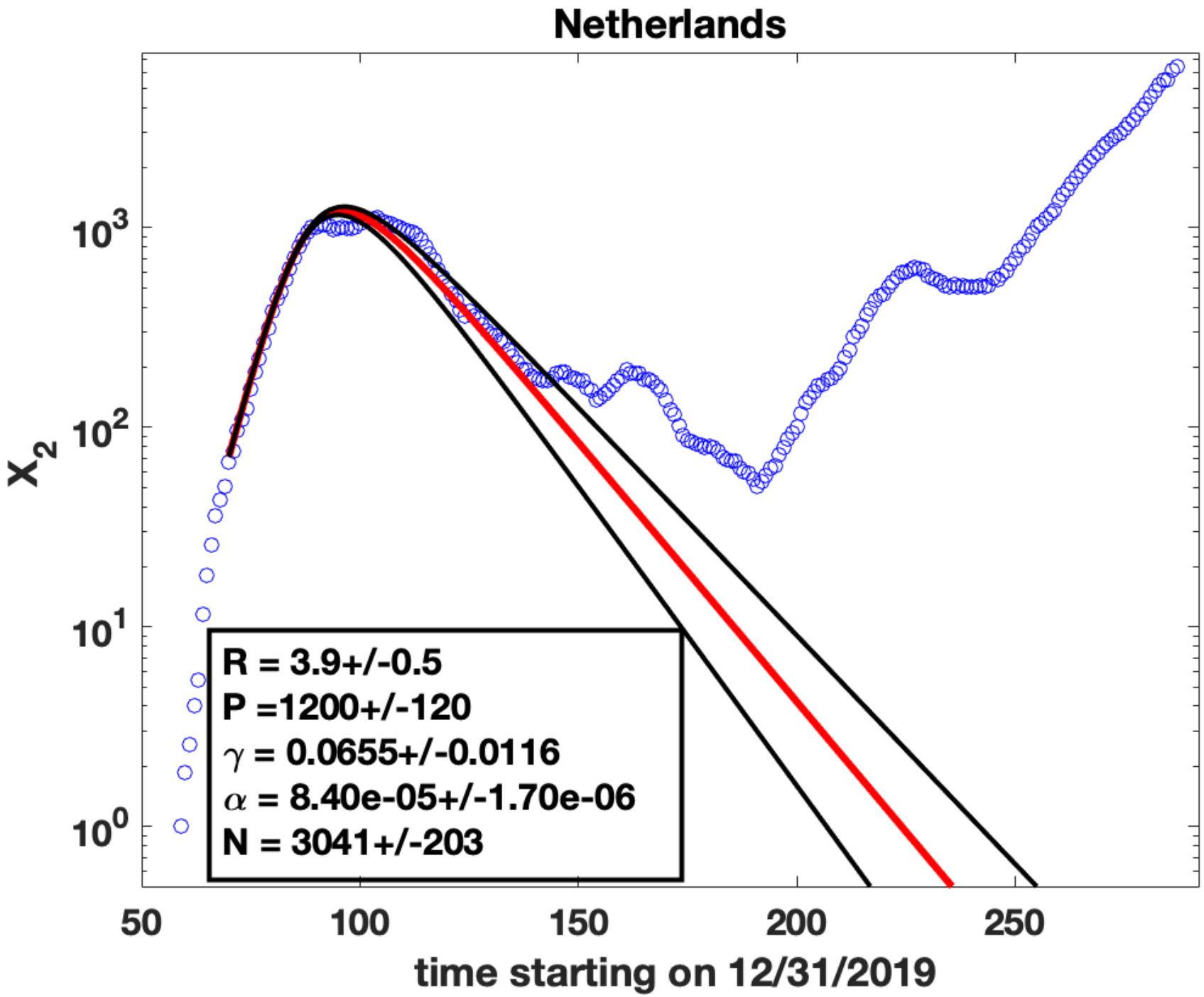

Netherlands
10^3
10^2
10^1
10^0
X_2
R = 3.9+/-0.5
P =1200+/-120
γ = 0.0655+/-0.0116
α = 8.40e-05+/-1.70e-06
N = 3041+/-203
50
100
150
200
250
time starting on 12/31/2019

Fig. 2b

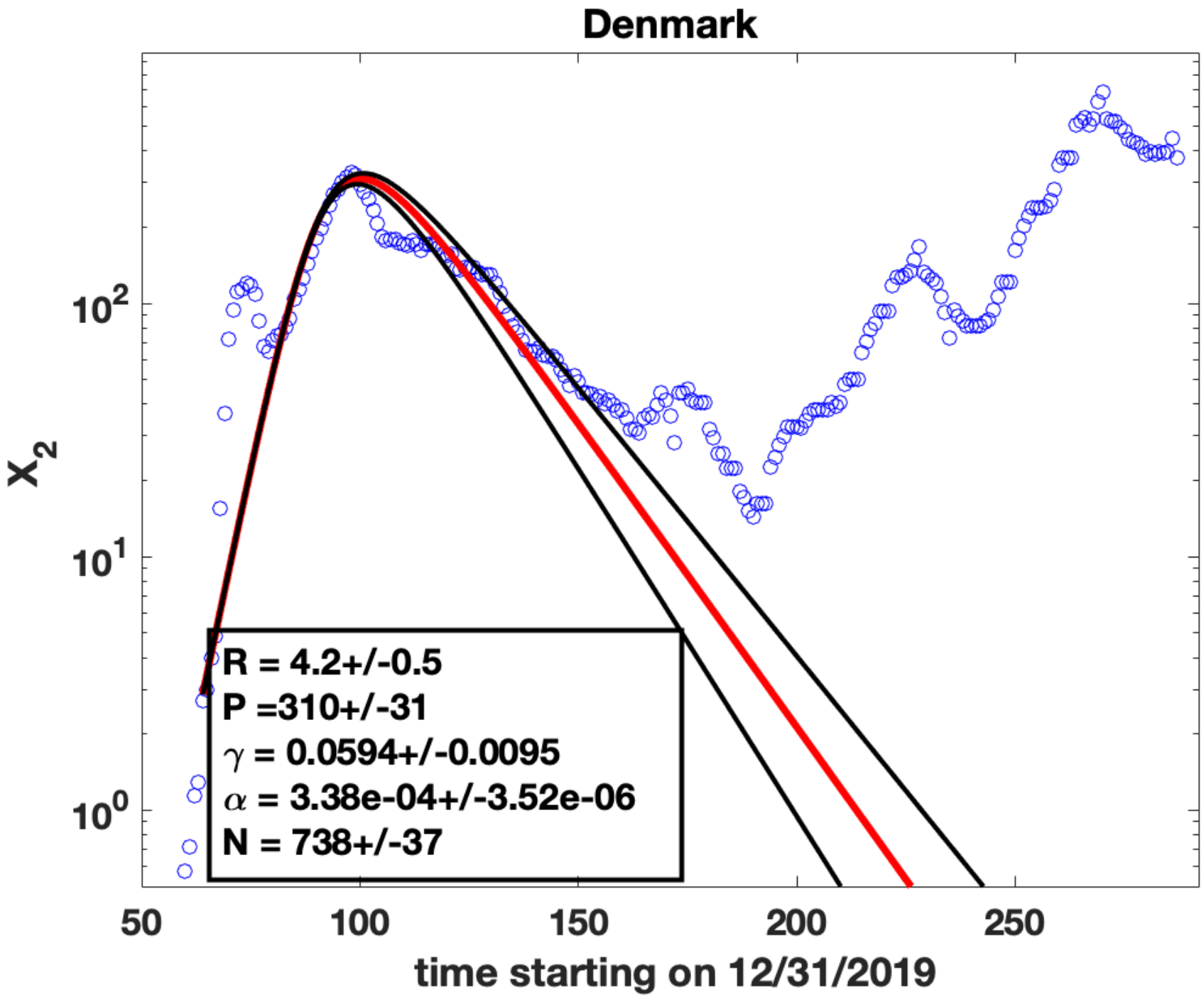

Denmark
10^2
10^1
10^0
X_2
R = 4.2+/-0.5
P =310+/-31
γ = 0.0594+/-0.0095
α = 3.38e-04+/-3.52e-06
N = 738+/-37
50
100
150
200
250
time starting on 12/31/2019

Fig. 2c

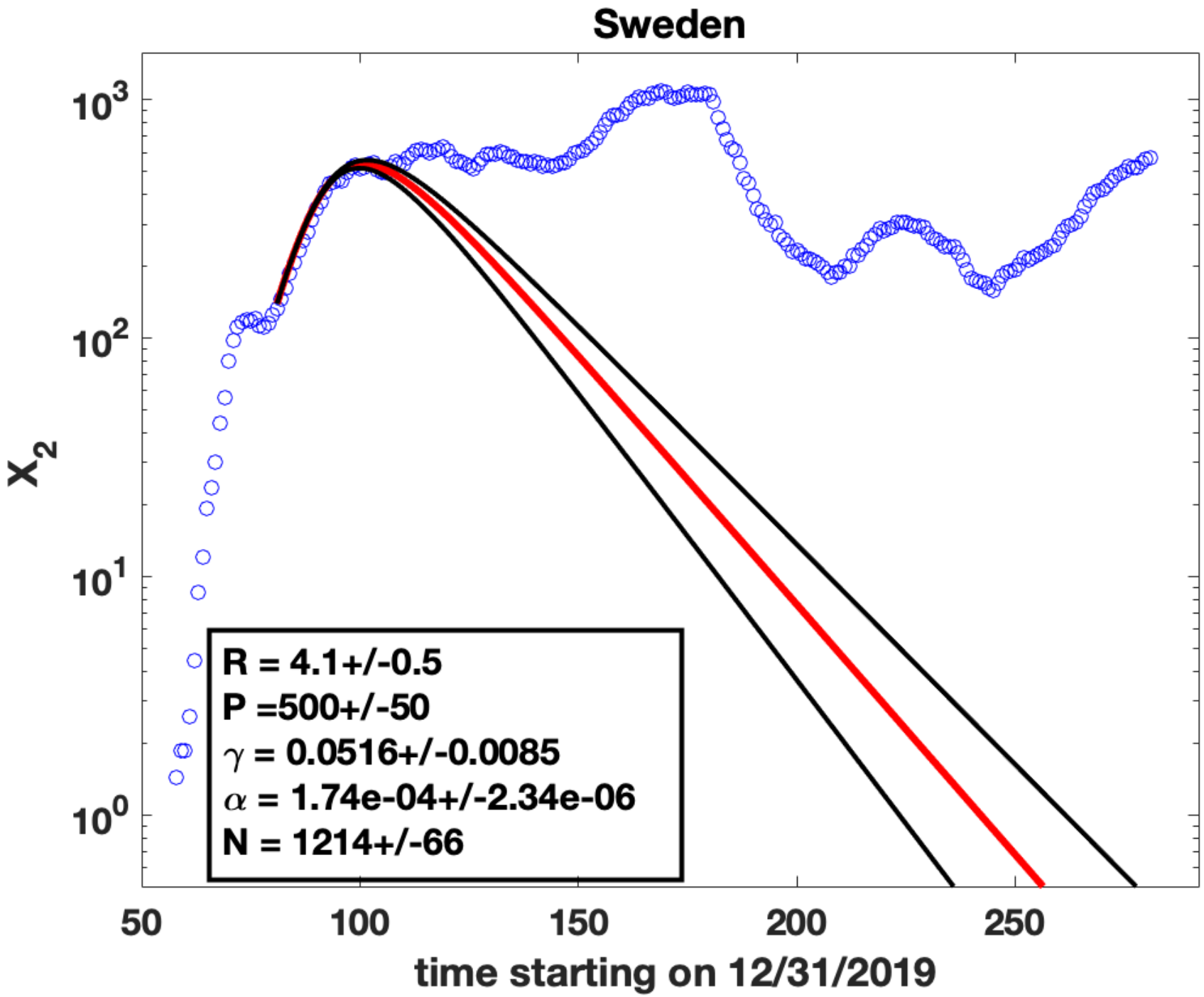

Sweden
10^3
10^2
10^1
10^0
X_2
R = 4.1+/-0.5
P =500+/-50
γ = 0.0516+/-0.0085
α = 1.74e-04+/-2.34e-06
N = 1214+/-66
50
100
150
200
250
time starting on 12/31/2019

Fig. 2d

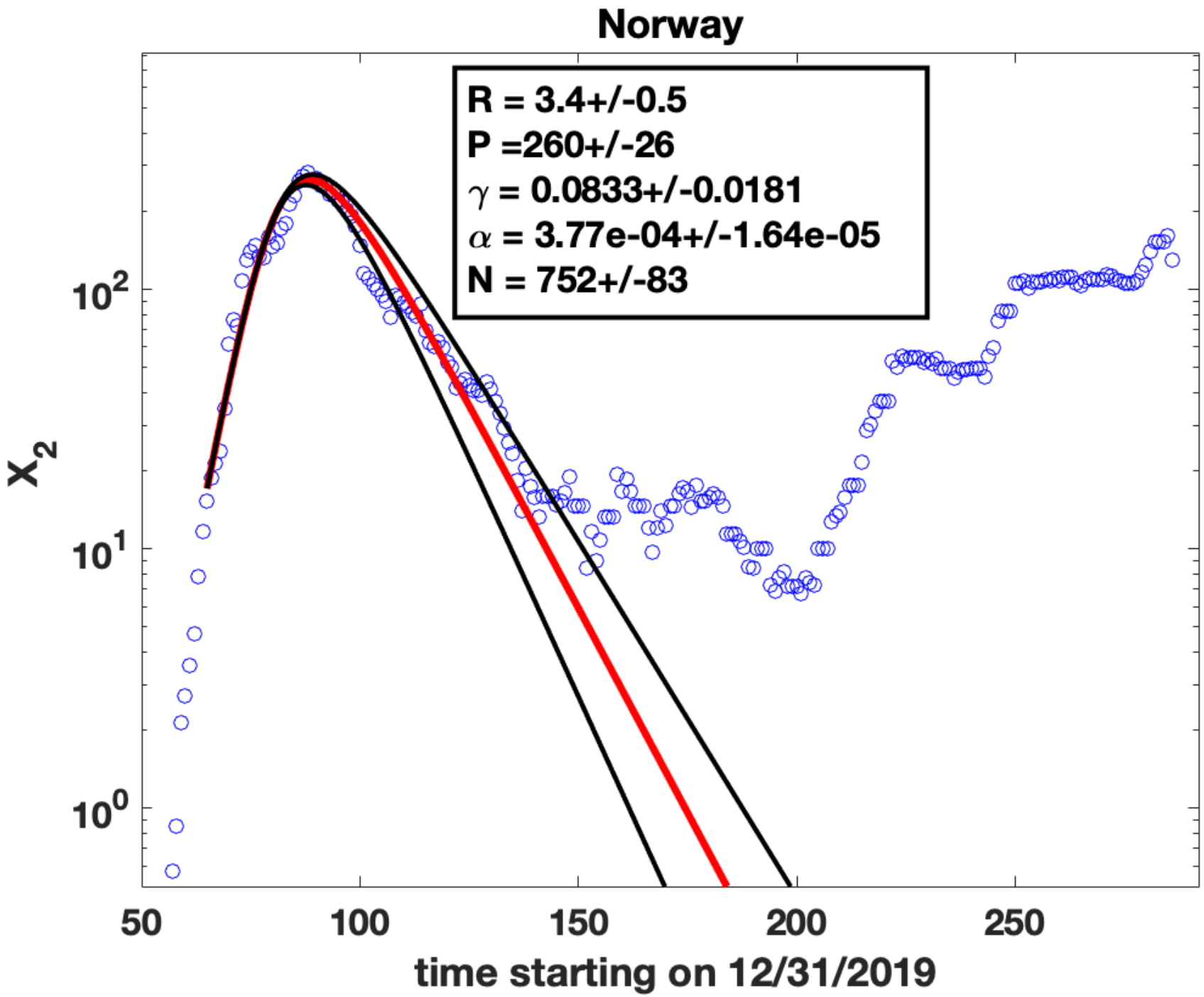

Norway
R = 3.4+/-0.5
P =260+/-26
γ = 0.0833+/-0.0181
α = 3.77e-04+/-1.64e-05
N = 752+/-83
10^2
10^1
10^0
X_2
50
100
150
200
250
time starting on 12/31/2019

Fig. 2e

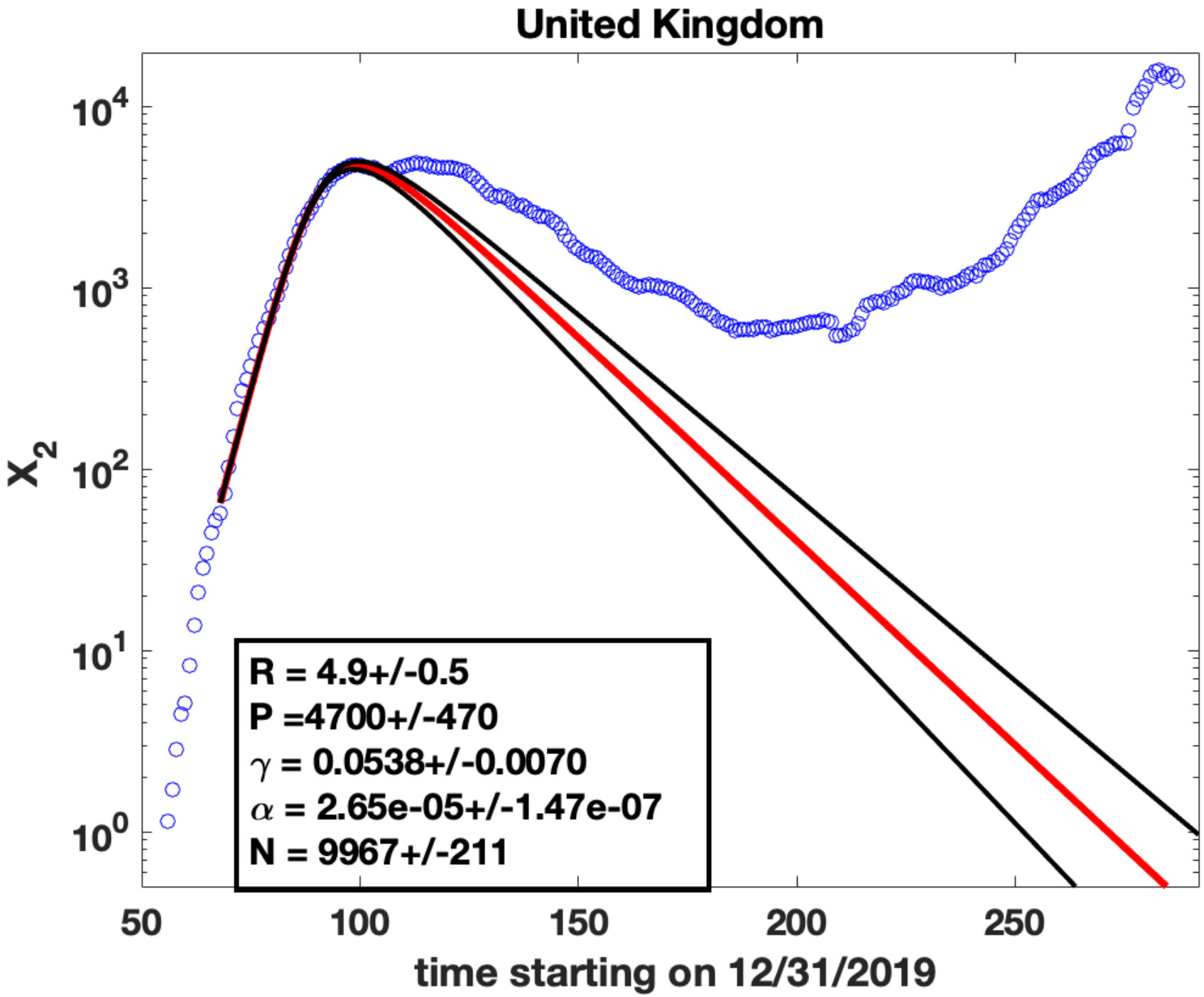

United Kingdom
10^4
10^3
10^2
10^1
10^0
X_2
R = 4.9+/-0.5
P =4700+/-470
γ = 0.0538+/-0.0070
α = 2.65e-05+/-1.47e-07
N = 9967+/-211
50
100
150
200
250
time starting on 12/31/2019

Fig. 2f

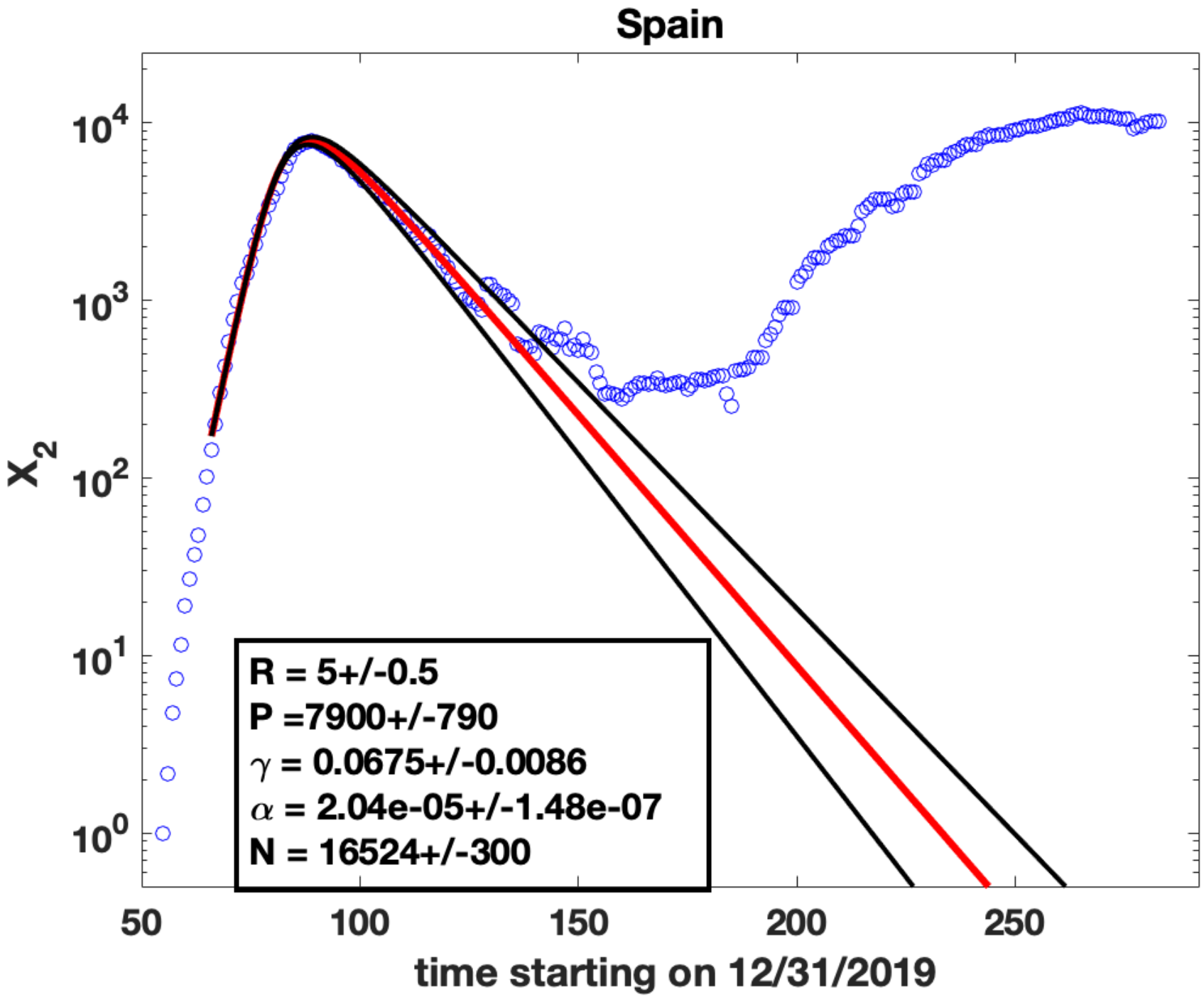

Spain
10^4
10^3
10^2
10^1
10^0
X_2
R = 5+/-0.5
P =7900+/-790
γ = 0.0675+/-0.0086
α = 2.04e-05+/-1.48e-07
N = 16524+/-300
50
100
150
200
250
time starting on 12/31/2019

Fig. 2g

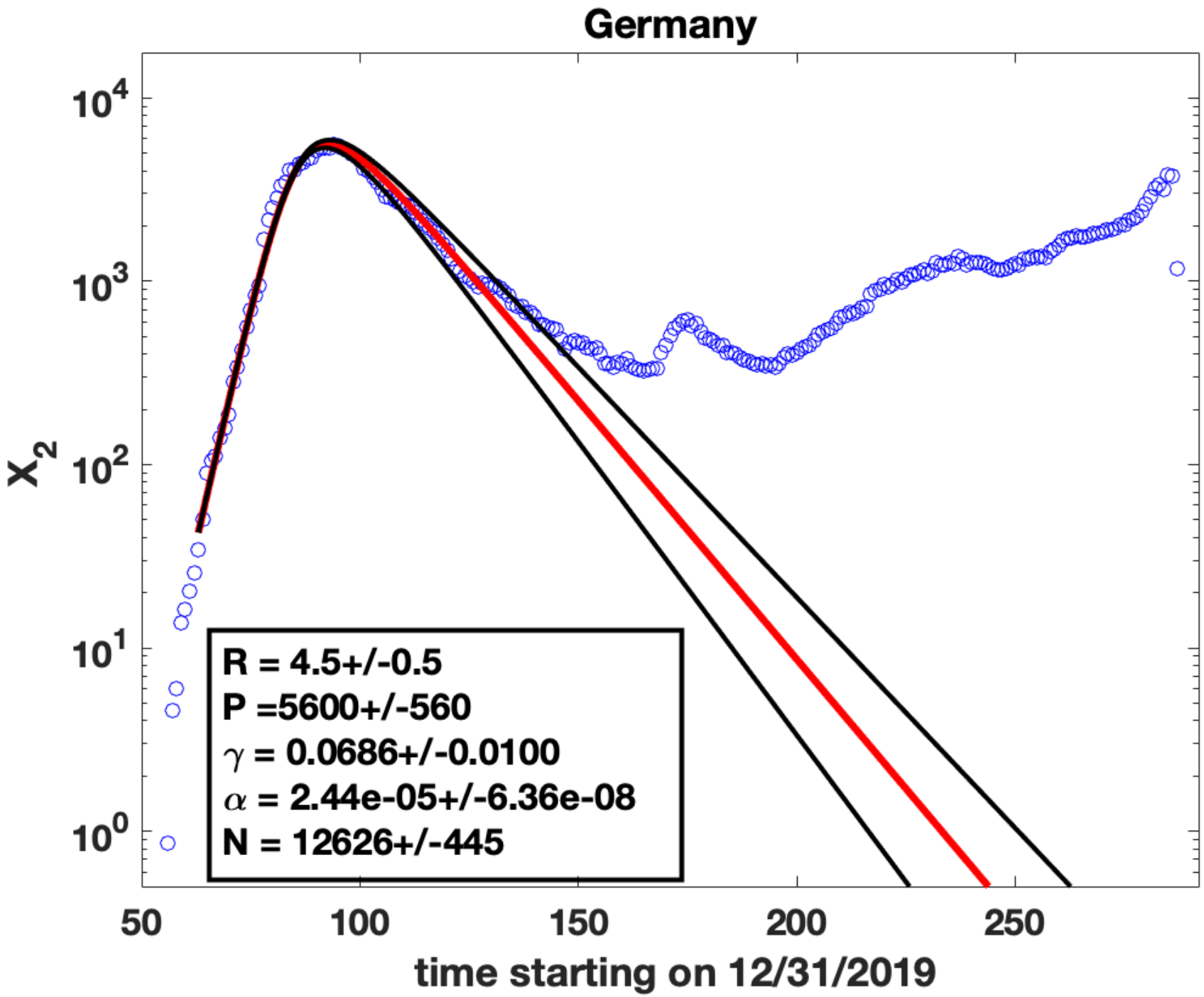

Germany
10^4
10^3
10^2
10^1
10^0
X_2
R = 4.5+/-0.5
P =5600+/-560
γ = 0.0686+/-0.0100
α = 2.44e-05+/-6.36e-08
N = 12626+/-445
50
100
150
200
250
time starting on 12/31/2019

Fig. 2h

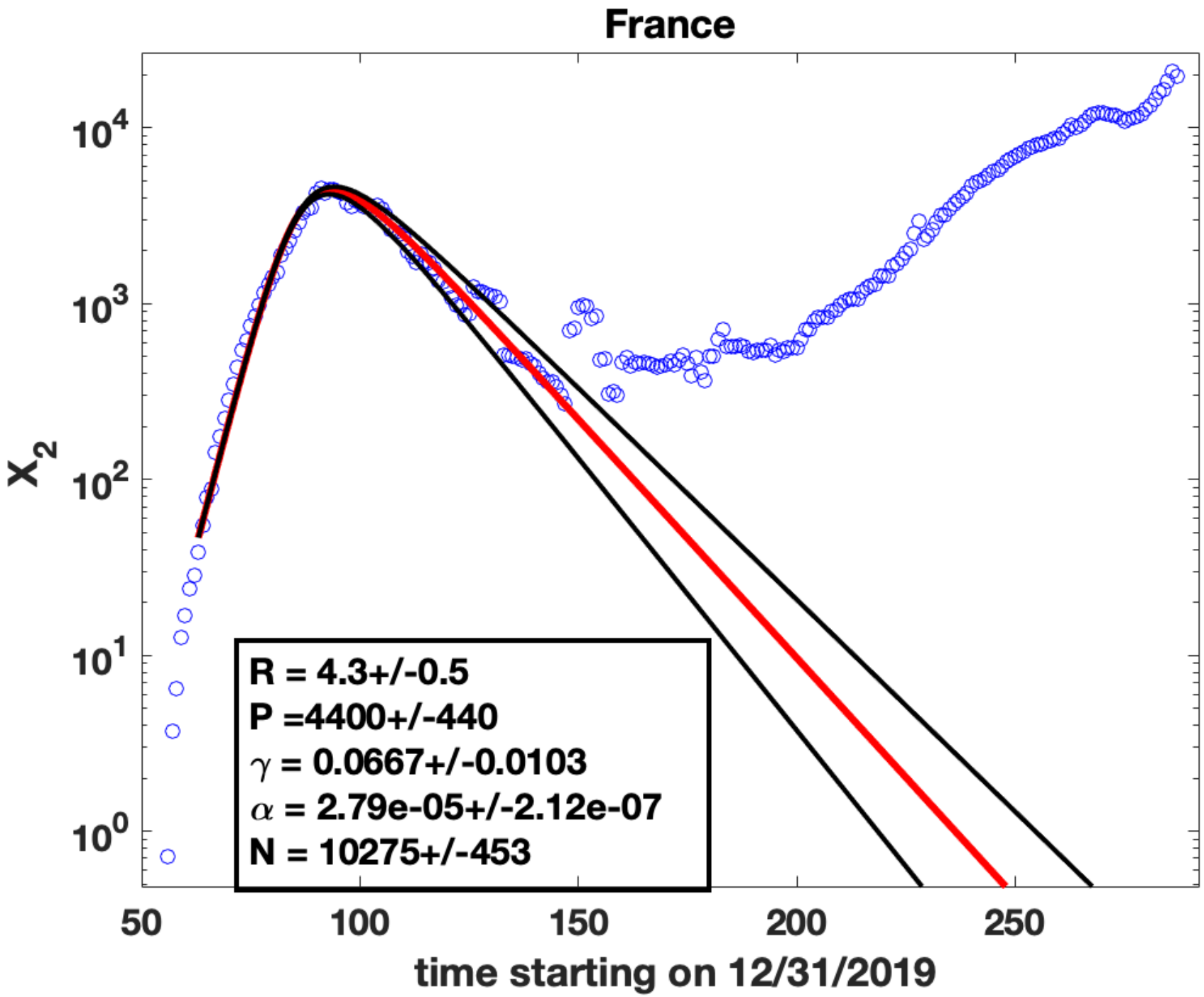

France
10^4
10^3
10^2
10^1
10^0
X_2
R = 4.3+/-0.5
P =4400+/-440
γ = 0.0667+/-0.0103
α = 2.79e-05+/-2.12e-07
N = 10275+/-453
50
100
150
200
250
time starting on 12/31/2019

Fig. 2i

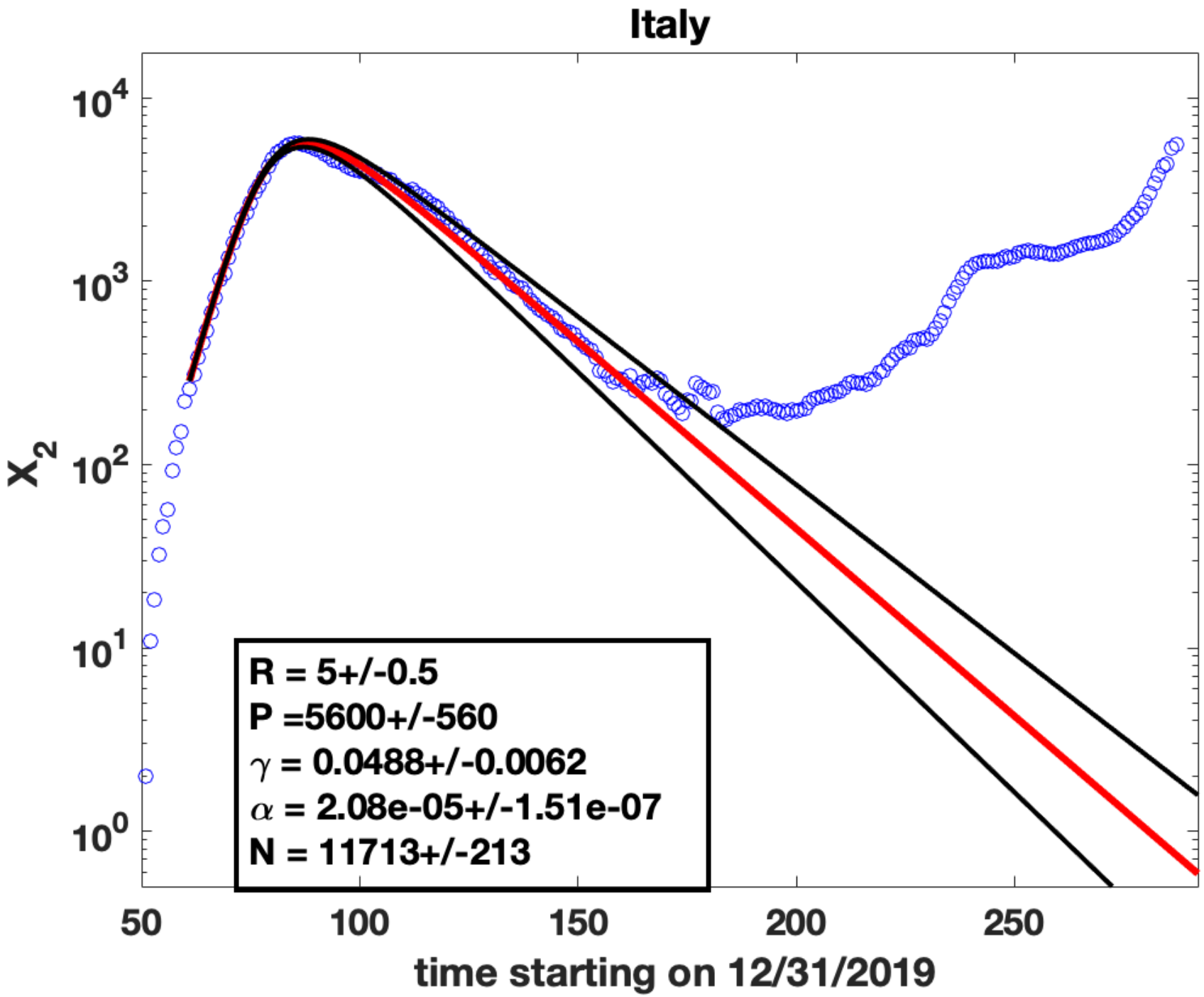

Italy
10^4
10^3
10^2
10^1
10^0
X_2
R = 5+/-0.5
P =5600+/-560
γ = 0.0488+/-0.0062
α = 2.08e-05+/-1.51e-07
N = 11713+/-213
50
100
150
200
250
time starting on 12/31/2019

Fig. 3a

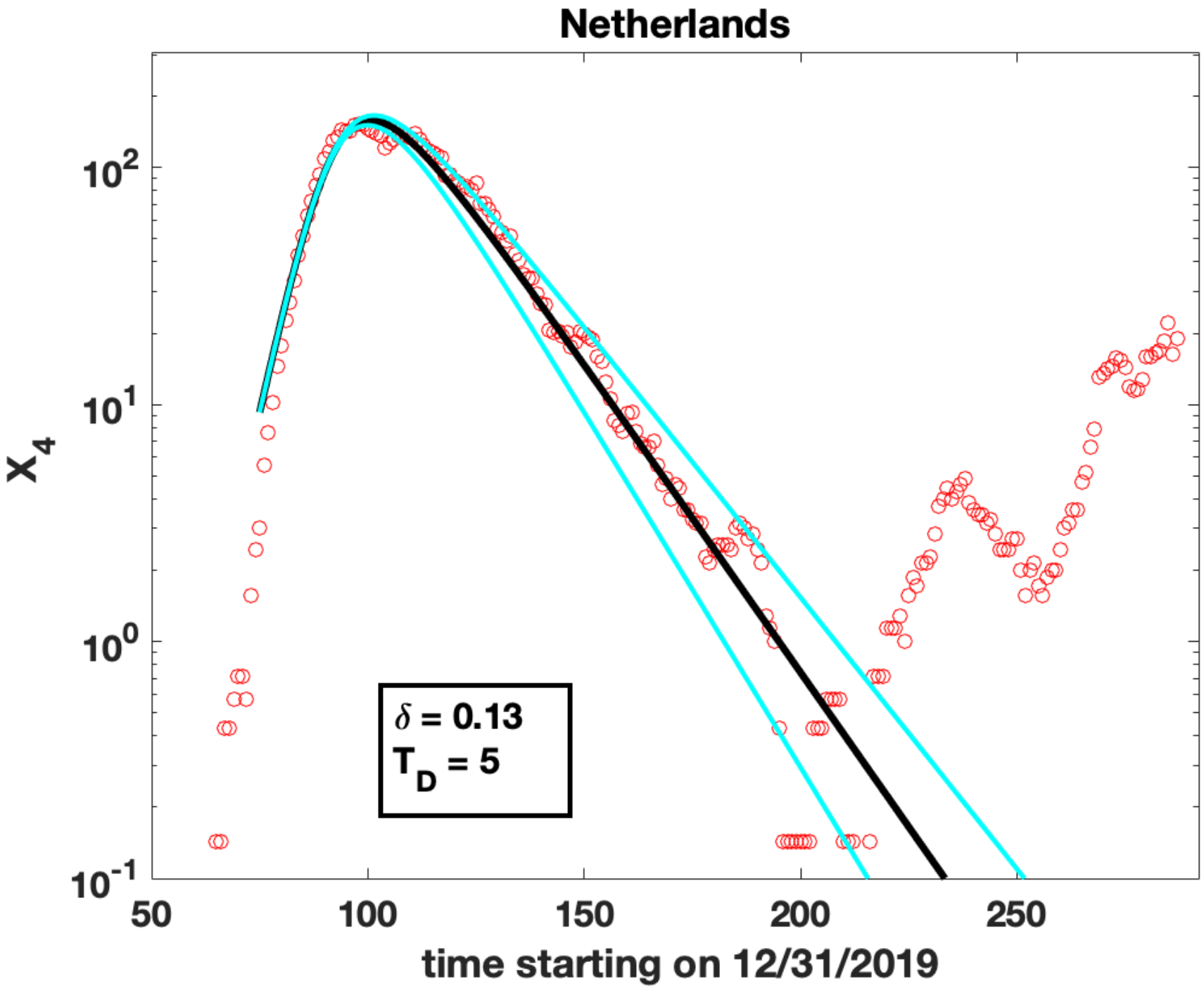

Netherlands
10^2
10^1
10^0
10^{-1}
X_4
50
100
150
200
250
time starting on 12/31/2019
δ = 0.13
T_D = 5

Fig. 3b

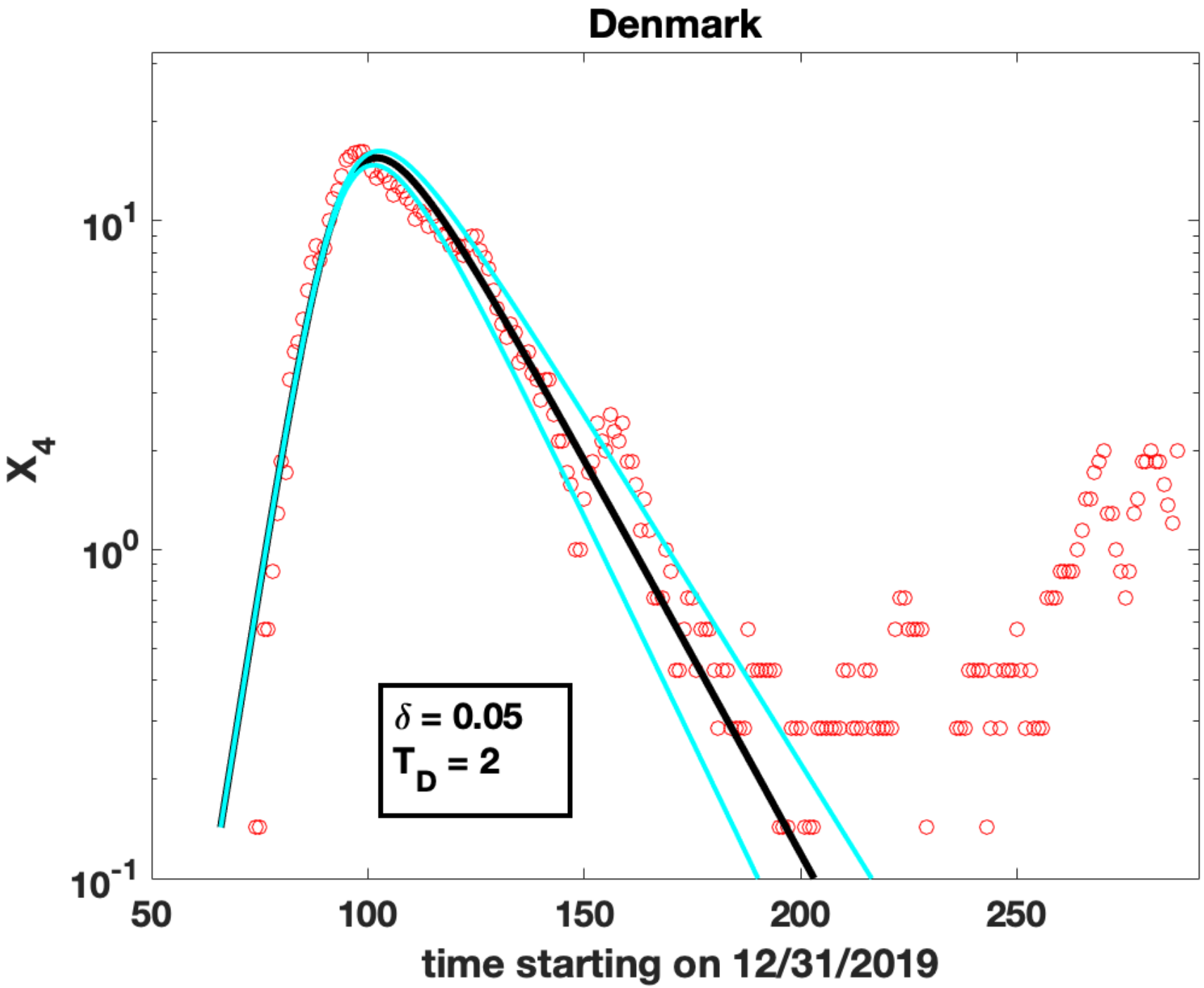

Denmark
10^1
10^0
10^{-1}
X_4
50
100
150
200
250
time starting on 12/31/2019
δ = 0.05
T_D = 2

Fig. 3c

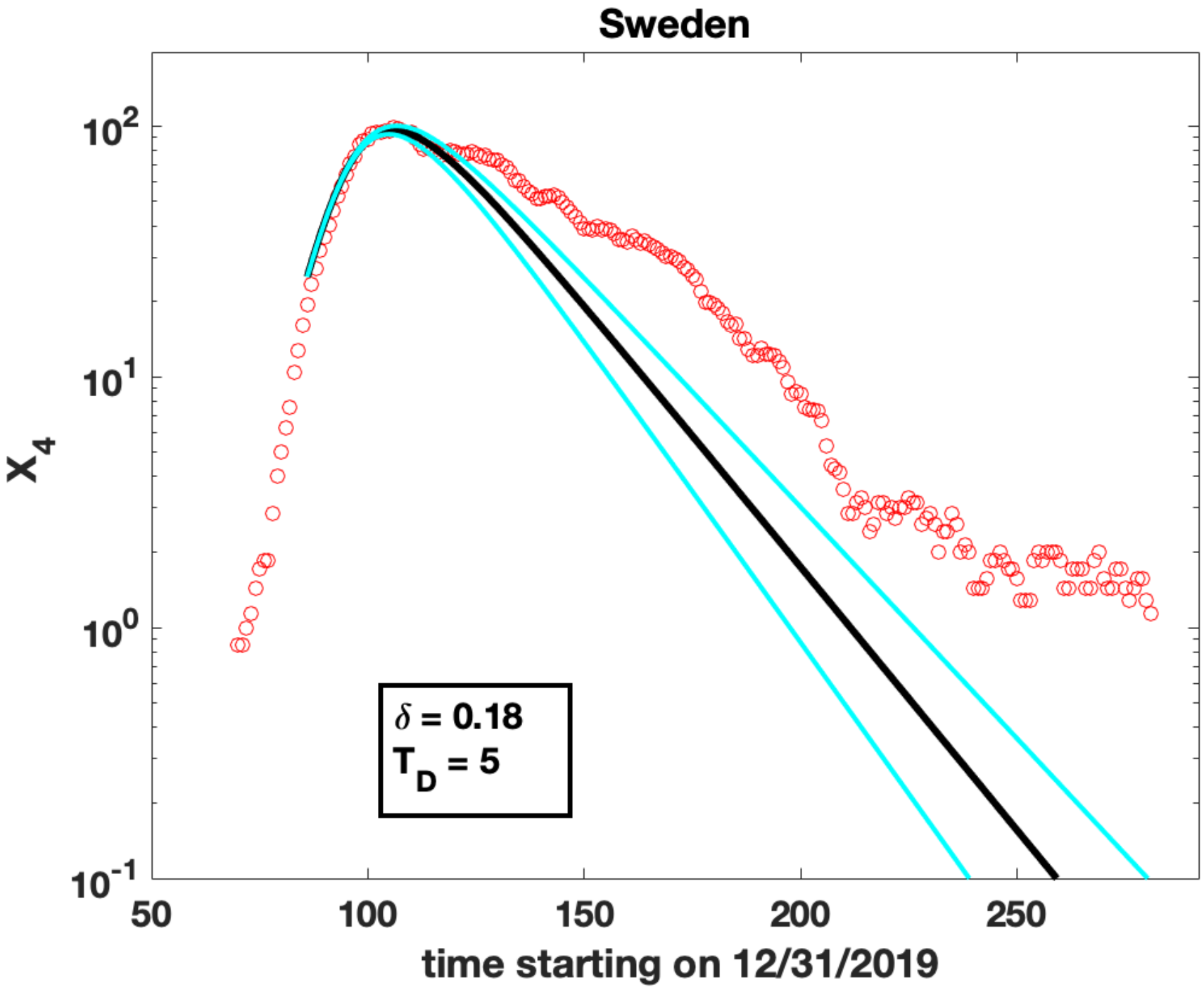

Sweden
10^2
10^1
10^0
10^{-1}
X_4
50
100
150
200
250
time starting on 12/31/2019
δ = 0.18
T_D = 5

Fig. 3d

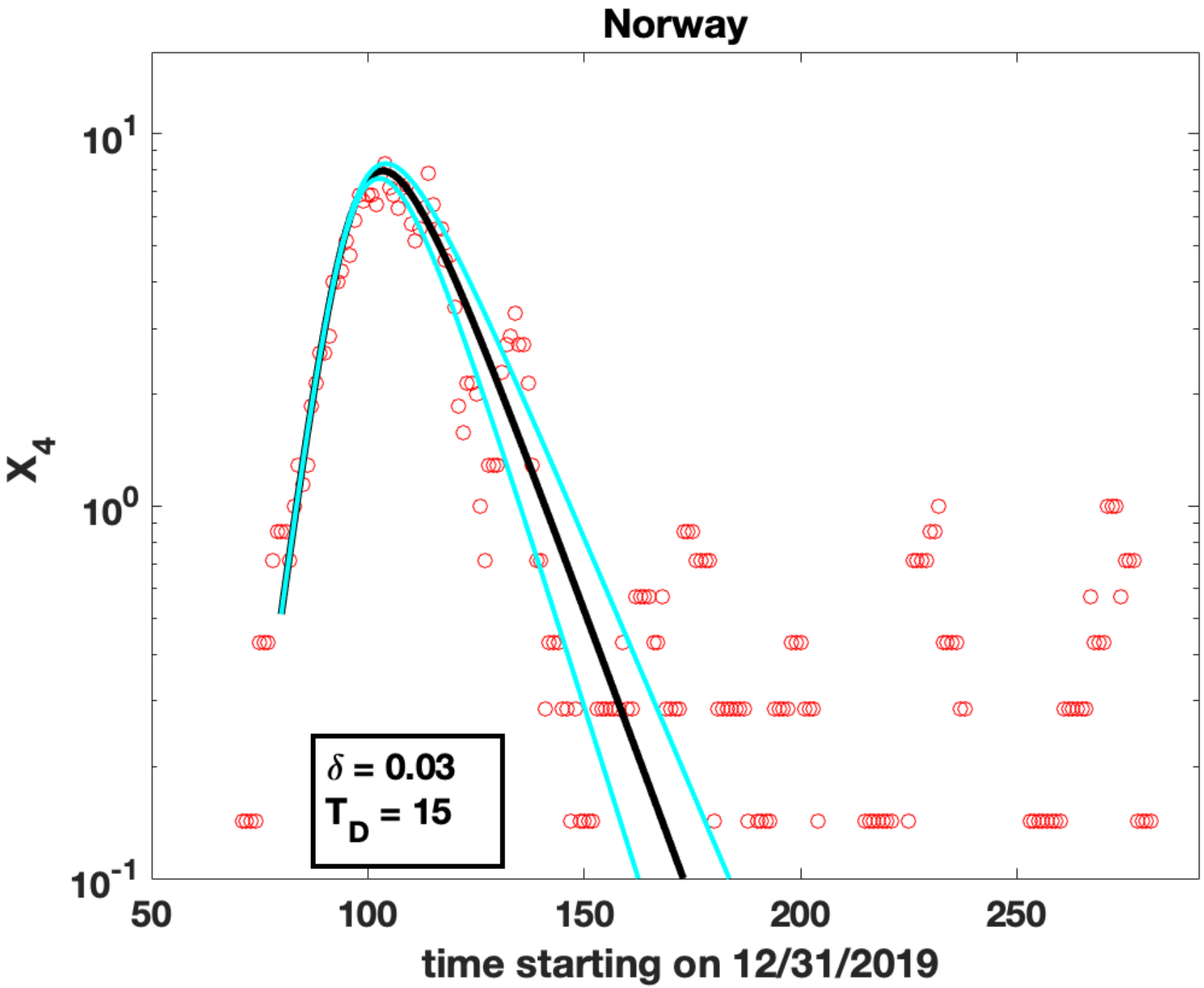

Norway
10^1
10^0
10^{-1}
X_4
50
100
150
200
250
time starting on 12/31/2019
δ = 0.03
T_D = 15

Fig. 3e

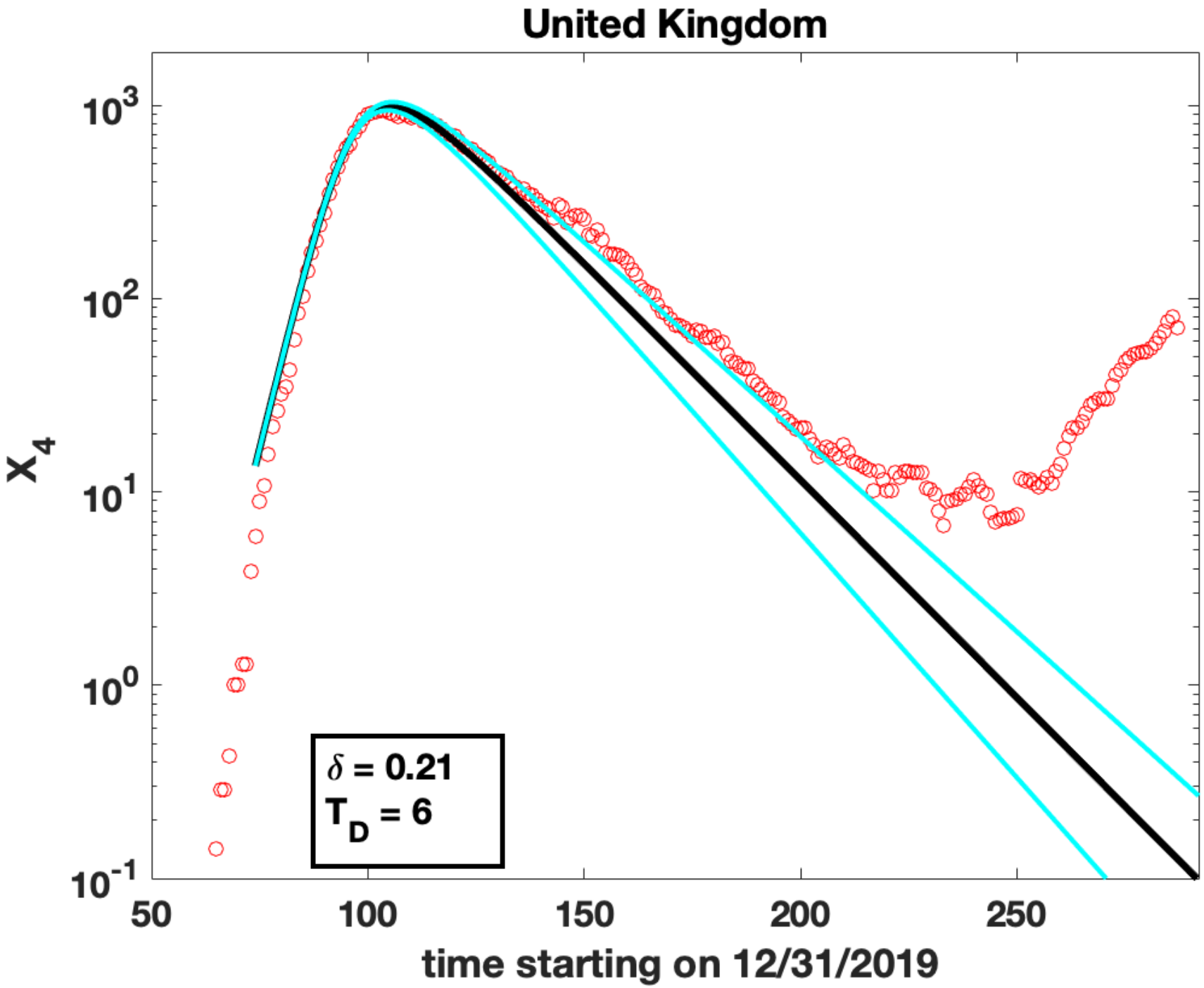

United Kingdom
10^3
10^2
10^1
10^0
10^{-1}
X_4
50
100
150
200
250
time starting on 12/31/2019
δ = 0.21
T_D = 6

Fig. 3f

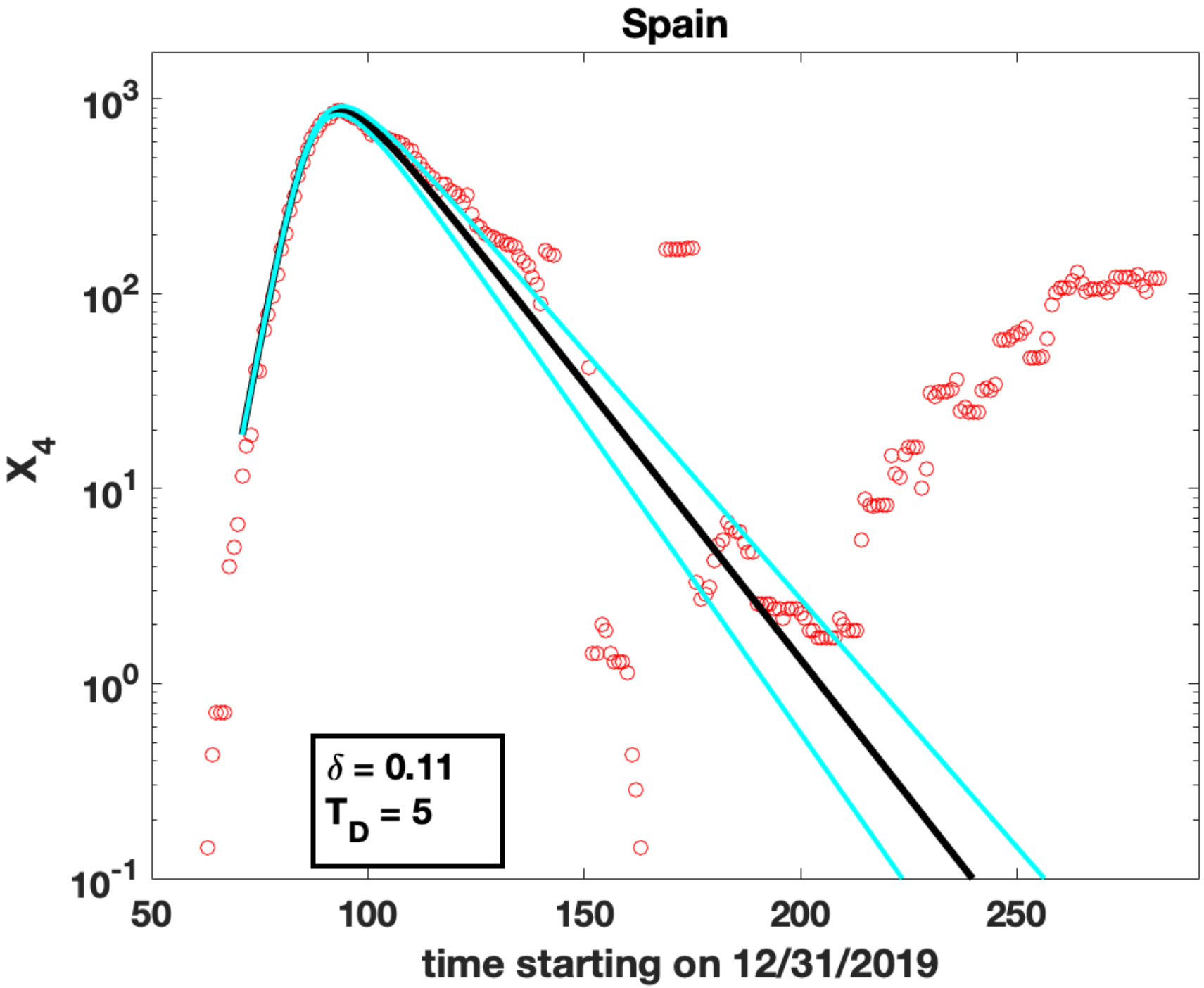

Spain
10^3
10^2
10^1
10^0
10^{-1}
X_4
50
100
150
200
250
time starting on 12/31/2019
δ = 0.11
T_D = 5

Fig. 3g

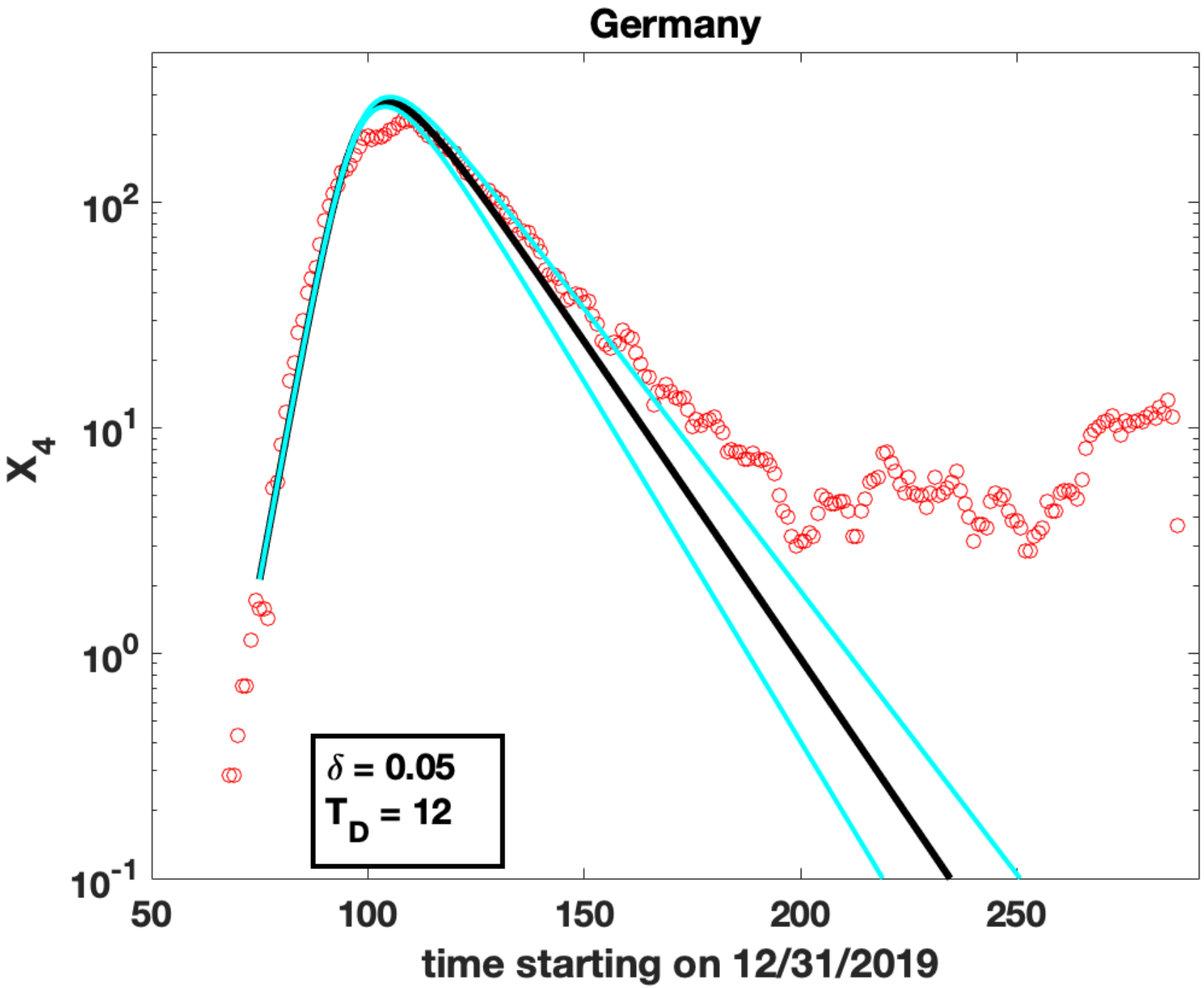

Germany
10^2
10^1
10^0
10^{-1}
X_4
50
100
150
200
250
time starting on 12/31/2019
δ = 0.05
T_D = 12

Fig. 3h

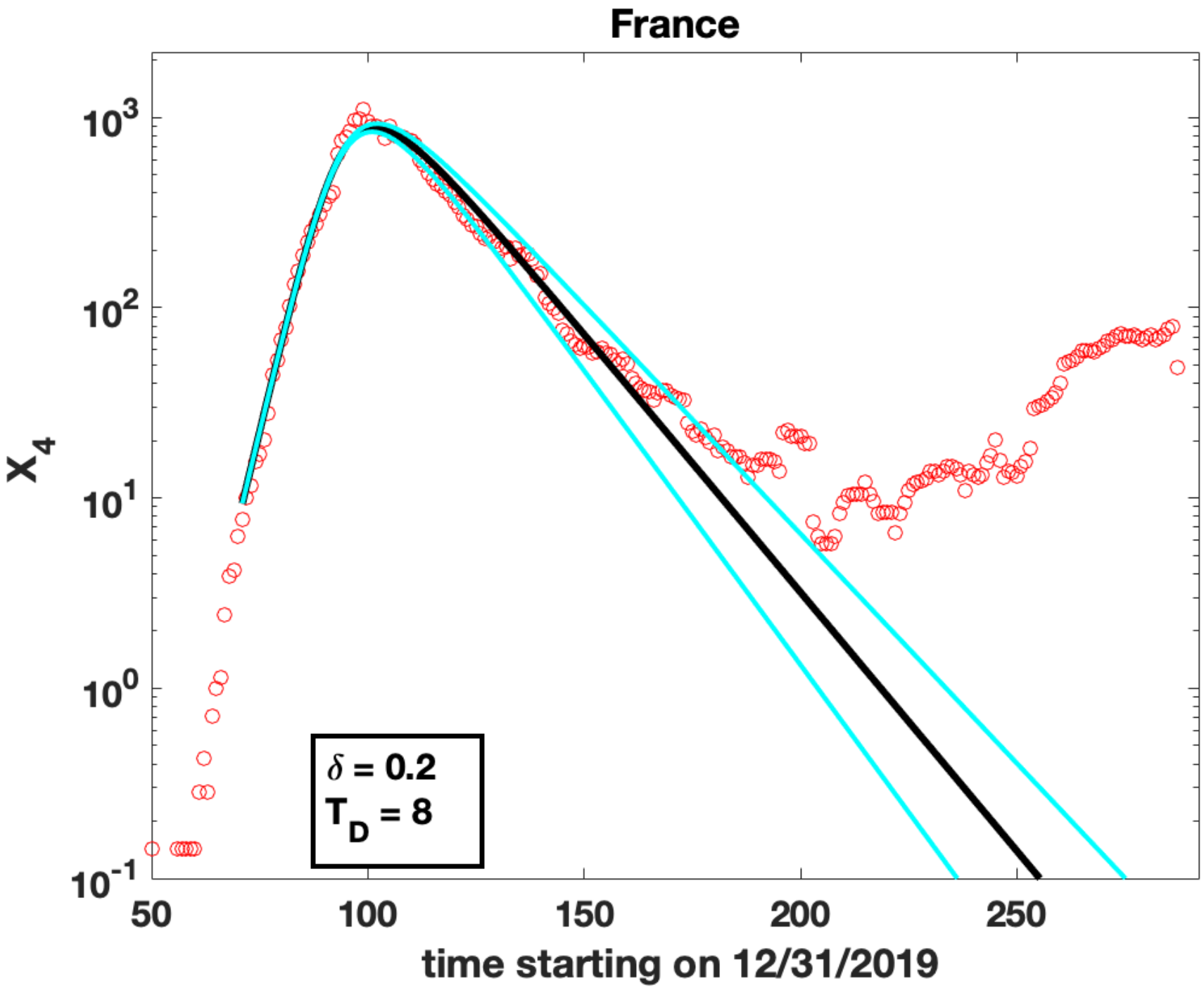

France
10^3
10^2
10^1
10^0
10^{-1}
X_4
50
100
150
200
250
time starting on 12/31/2019
δ = 0.2
T_D = 8

Fig. 3i

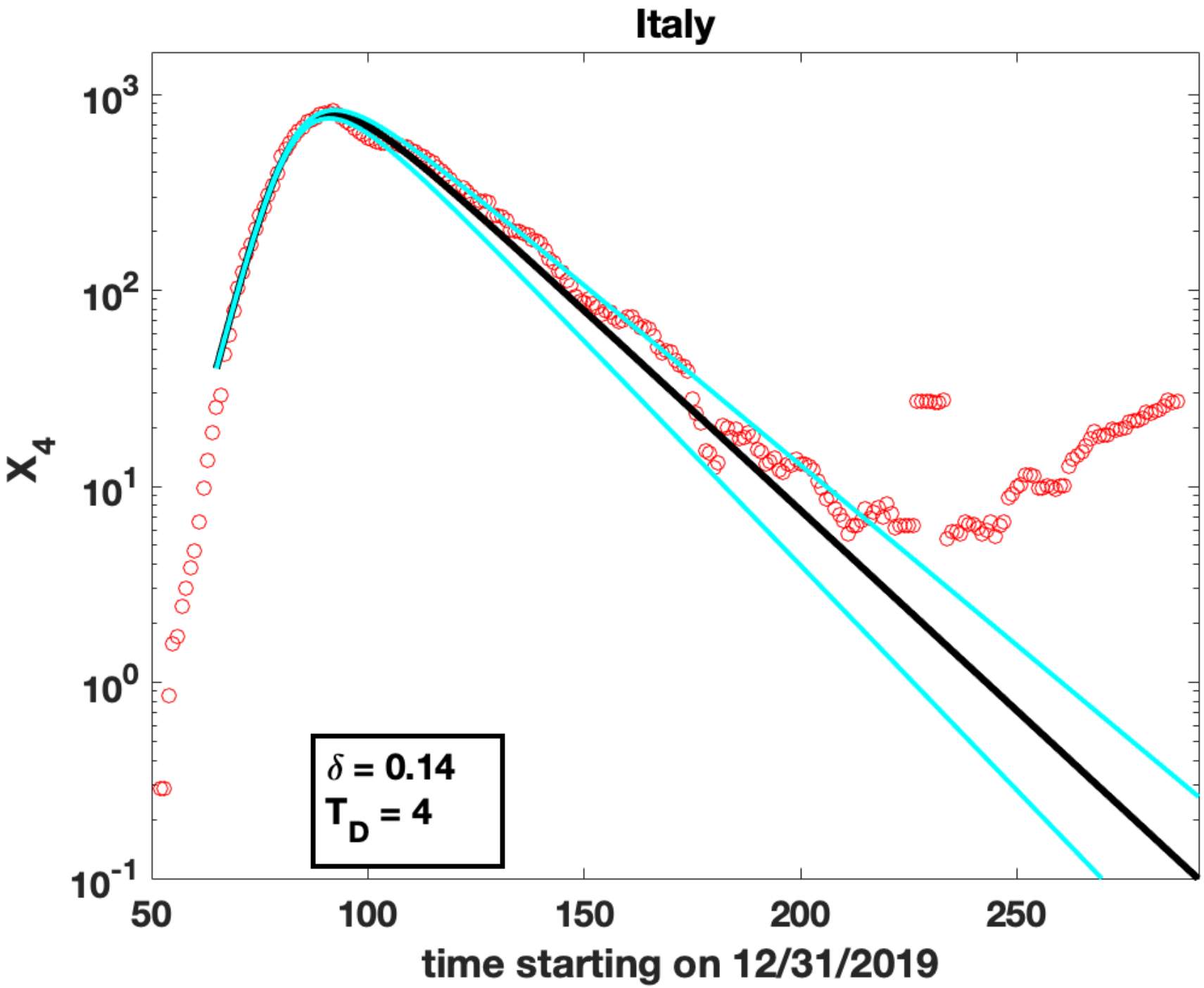

Italy
10^3
10^2
10^1
10^0
10^{-1}
X_4
50
100
150
200
250
time starting on 12/31/2019
δ = 0.14
T_D = 4

Fig. 4a

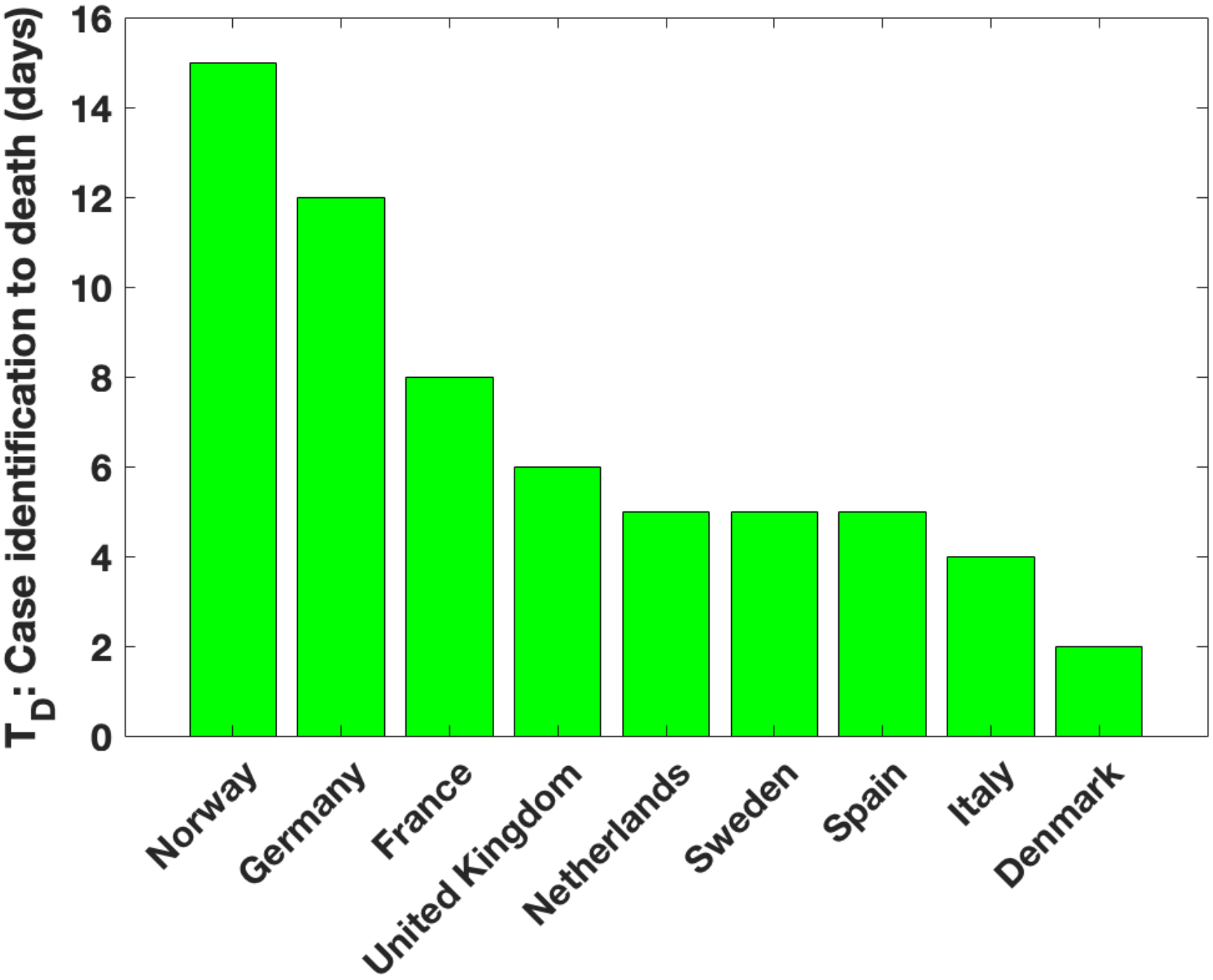

T_D: Case identification to death (days)
0
2
4
6
8
10
12
14
16
Norway
Germany
France
United Kingdom
Netherlands
Sweden
Spain
Italy
Denmark

Fig. 4b

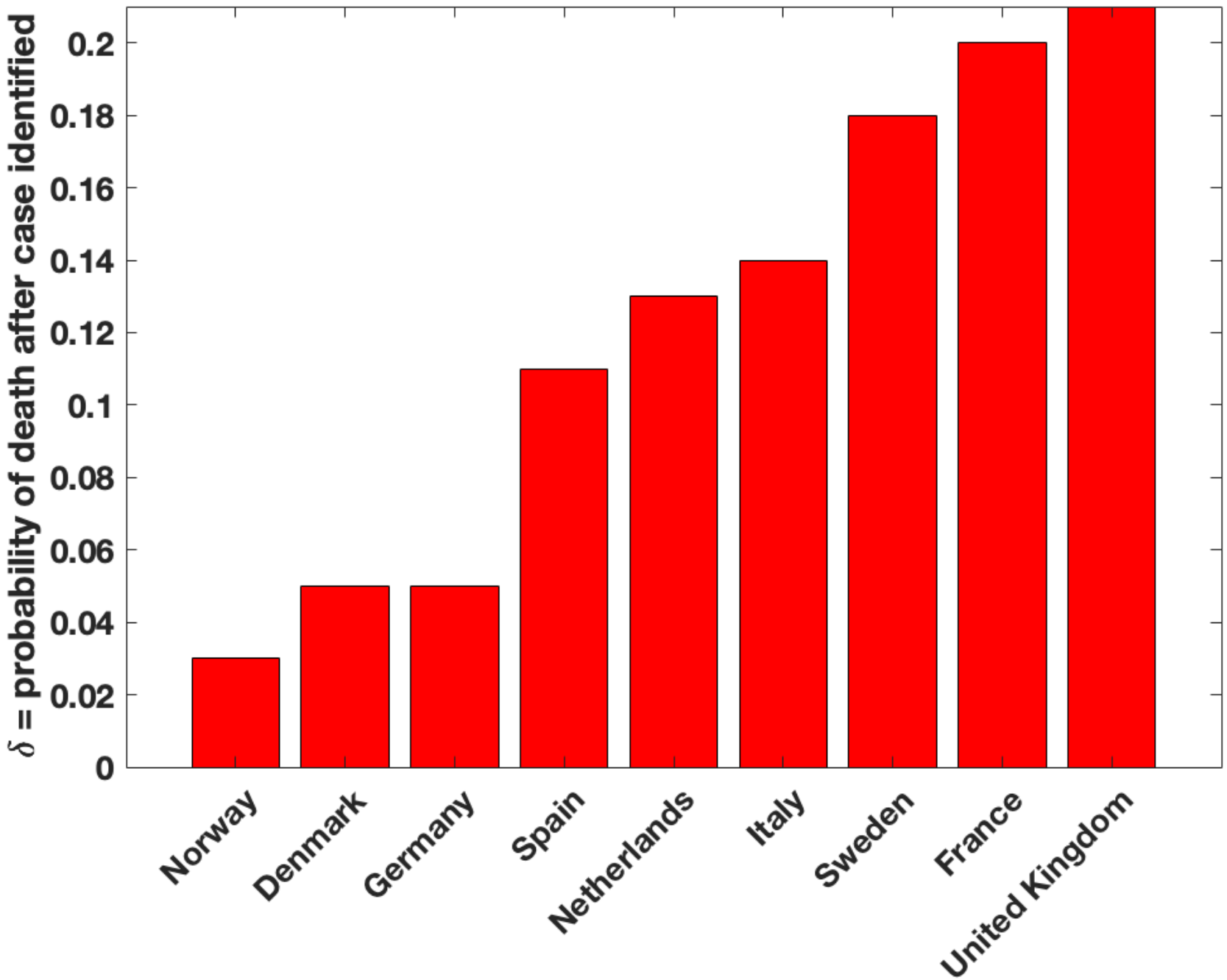

δ = probability of death after case identified
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
0.18
0.2
Norway
Denmark
Germany
Spain
Netherlands
Italy
Sweden
France
United Kingdom

Fig. 4c

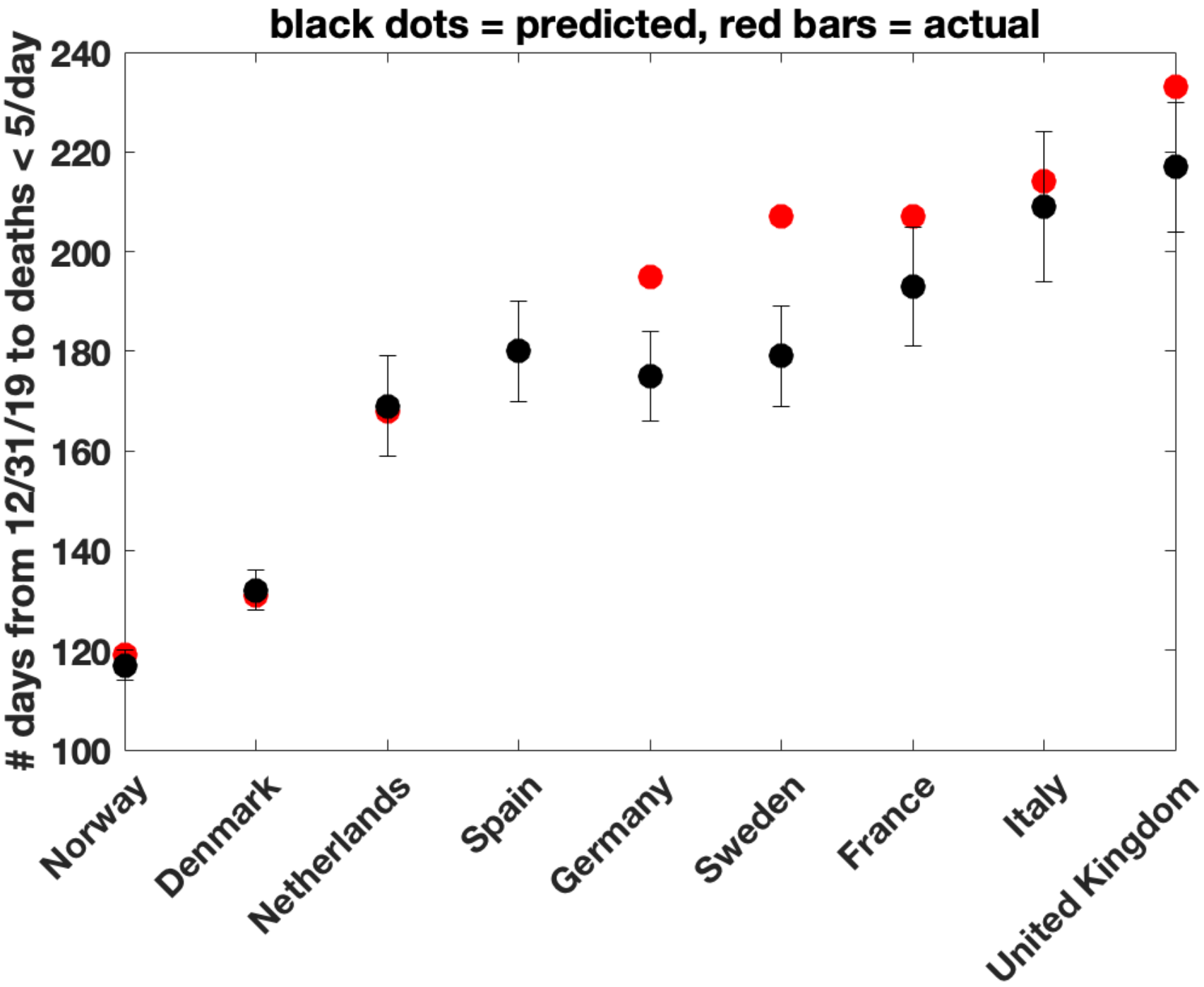

black dots = predicted, red bars = actual
days from 12/31/19 to deaths < 5/day
240
220
200
180
160
140
120
100
Norway
Denmark
Netherlands
Spain
Germany
Sweden
France
Italy
United Kingdom

Fig. 4d

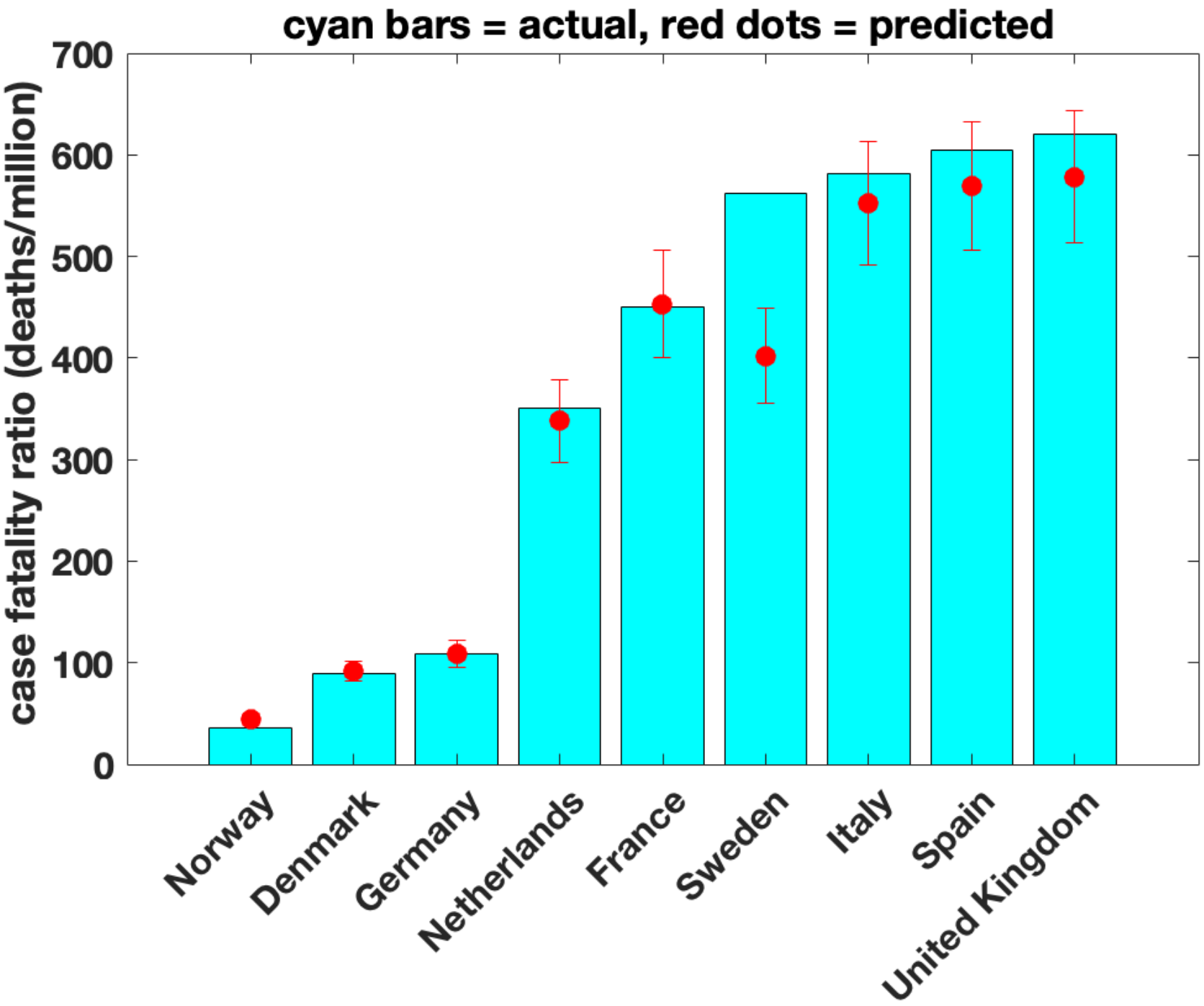

cyan bars = actual, red dots = predicted
case fatality ratio (deaths/million)
0
100
200
300
400
500
600
700
Norway
Denmark
Germany
Netherlands
France
Sweden
Italy
Spain
United Kingdom

Fig. 4e

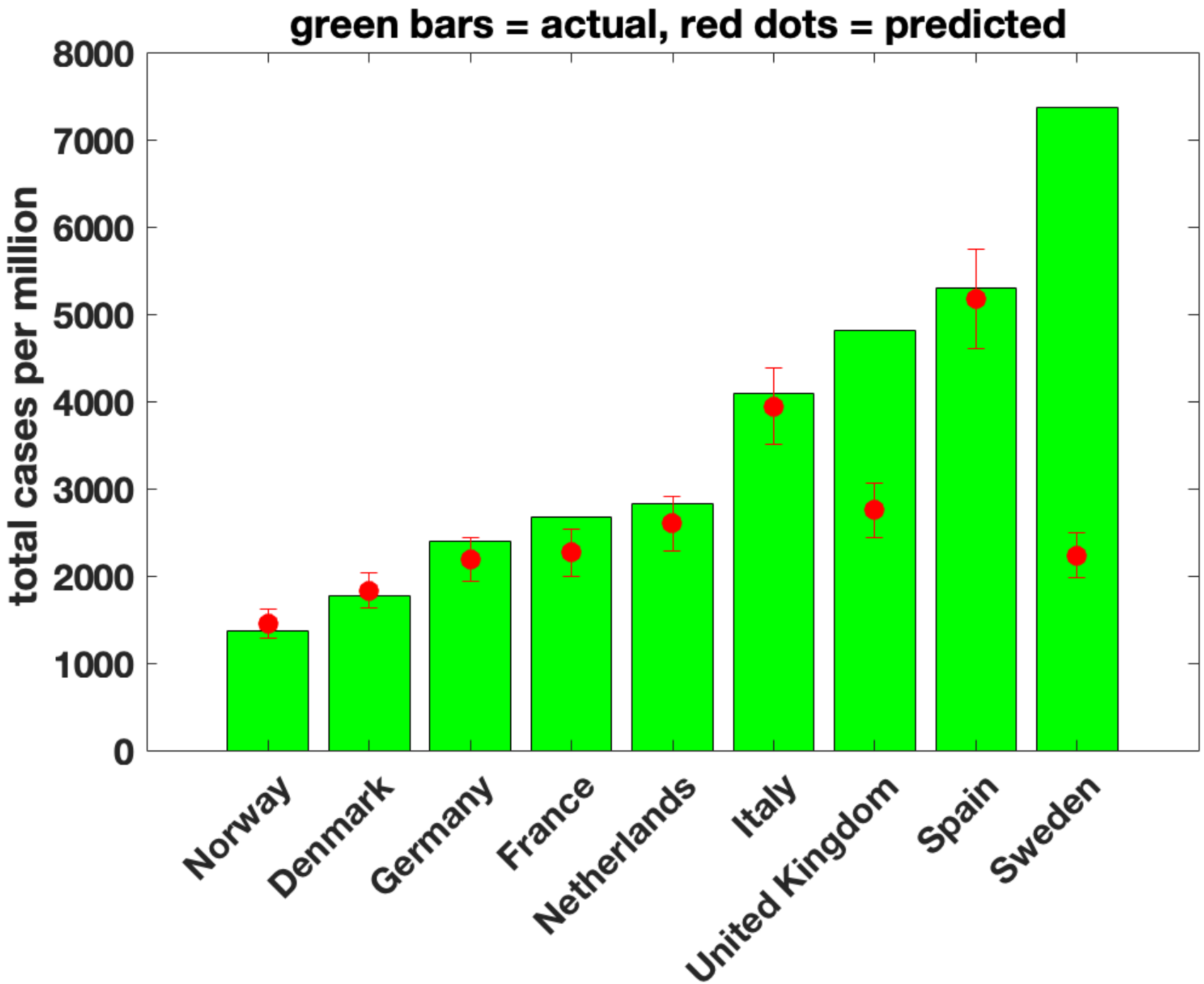
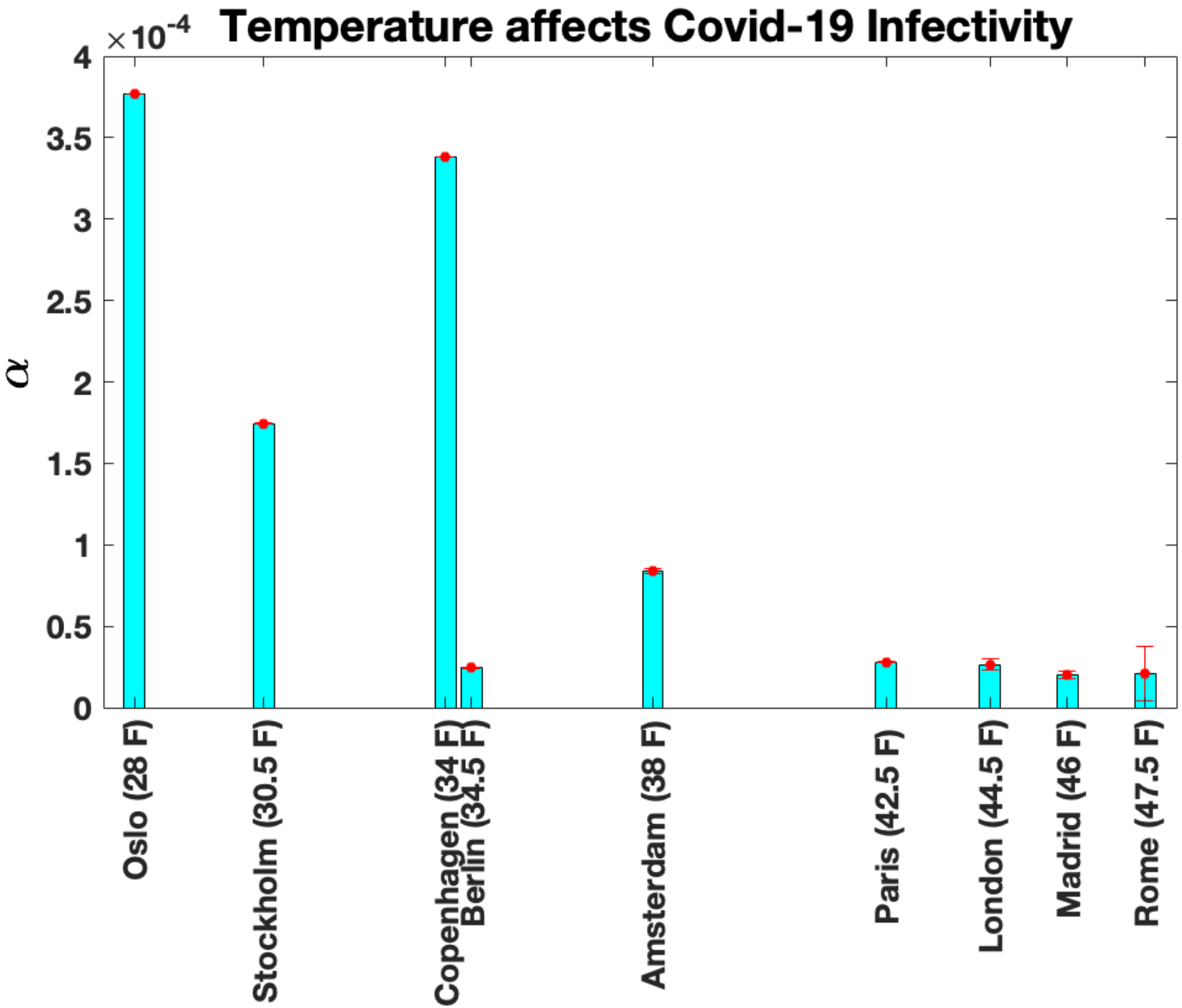

green bars = actual, red dots = predicted
total cases per million
8000
7000
6000
5000
4000
3000
2000
1000
0
Norway
Denmark
Germany
France
Netherlands
Italy
United Kingdom
Spain
Sweden

Fig. 4f


Temperature affects Covid-19 Infectivity
×10^-4
4
3.5
3
2.5
2
1.5
1
0.5
0
α
Oslo (28 F)
Stockholm (30.5 F)
Copenhagen (34 F)
Berlin (34.5 F)
Amsterdam (38 F)
Paris (42.5 F)
London (44.5 F)
Madrid (46 F)
Rome (47.5 F)

medRxiv preprint doi: https://doi.org/10.1101/2020.05.26.20114058; this version posted October 15, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

	NETHERLANDS	DENMARK	SWEDEN	NORWAY	UK	SPAIN	GERMANY	FRANCE	ITALY
Mean Latitude	52.133	56.26	60.13	60.17	52.36	40.46	51.17	46.23	41.87
Population Miillions	17.3	5.81	10.2	5.4	66.65	46.9	83	67	60.4
Major City	Amsterdam	Copenhagen	Stockholm	Oslo	London	Madrid	Berlin	Paris	Rome
Average Temp Feb.2020 (deg F)	38	34	30.5	28	44.5	46	34.5	42.5	47.5
γ(R-1)	0.19	0.19	0.16	0.2	0.21	0.27	0.24	0.22	0.195
P	1,200	310	500	260	4,700	7,900	5600	4,400	5,600
N	3,041 +/-203	738 +/-37	1,214+/-66	752 +/-83	9,967 +/-211	16,524 +/-300	12,626 +/-445	10,275+/-453	11,713 +/-213
R	3.9 +/-0.5	4.2 +/-0.5	4.1 +/-0.5	3.4 +/-0.5	4.9 +/-0.5	5.0 +/-0.5	4.5 +/-0.5	4.3 +/-0.5	5.0 +/-0.50
αx 100,000	8.40 +/-0.17	33.80 +/-0.37	17.44 +/-0.23	37.7 +/-1.64	2.65 +/-0.02	2.04+/-0.02	2.44 +/-0.06	2.79+/-0.02	2.08 +/-0.02
γ	0.066 +/-0.012	0.059 +/-0.010	0.052 +/-0.009	0.083 +/-0.018	0.054 +/-0.007	0.068 +/-0.009	0.069 +/-0.01	0.067 +/-0.01	0.049 +/-0.006
δ	0.130	0.050	0.180	0.030	0.210	0.110	0.050	0.200	0.140
T_D	5	2	5	15	6	5	12	8	4
$T_L = 1/\gamma$	15.3	16.8	19.4	12.0	18.6	14.8	14.6	15.0	20.5
$T_R = 1/(N\alpha)$	3.9	4.0	4.7	3.5	3.8	3.0	3.2	3.5	4.1
Actual Max daily cases at first peak	1,120	328	532	282	4,751	7,902	5,596	4,482	5,644
Model Fitted Max daily cases at first peak	1,219+/-58	311+/-16	536+/-21	265+/-13	4,716+/-234	7,928+/-393	5,601+/-277	4,410+/-218	5659+/-274
Date when number of daily cases first peaked	4/5/20	4/9/20	4/10/20	3/28/20	4/8/20	3/29/20	4/2/20	4/3/20	3/28/20
N_{end}(P) = Predicted Number of days from 12/31/2019 to pandemic end defined as day Daily Deaths < 5 after first peak	169 +/- 10 days	132 +/- 4 days	179 +/- 10 days	117 +/- 3 days	217 +/- 13 days	180 +/- 10 days	175 +/- 9 days	193 +/- 12 days	209 +/- 15 days
T_{end}(P) = Predicted date pandemic ends = defined as day Daily Deaths < 5 after first peak	6/18/2020 +/- 10 days	5/11/2020 +/- 4 days	6/28/2020 +/- 10 days	4/26/2020 +/- 3 days	8/5/2020 +/- 13 days	6/29/2020 +/- 10 days	6/24/2020 +/- 9 days	7/12/2020 +/- 12 days	7/28/2020 +/- 15 days
N_{end}(A) = Actual number of days after first peak when Daily Deaths < 5	168	131	207	119	233	180	195	207	214
T_{end}(A) = Actual date after first peak when Daily Deaths < 5	6/17/20	5/10/20	7/26/20	4/28/20	8/21/20	6/29/20	7/14/20	7/19/20	8/2/20
Predicted Cases upto T_{end}(P)	45,015 +/-5,413	10,645+/-1,199	22,799 +/-2,657	7,844+/-881	183,393+/-20,530	242,705+/-26,975	181,375 +/-20,844	151,738 +/-17,726	238,163 +/-26,426
Actual Cases upto T_{end}(A)	48,921	10,306	75,228	7,382	320,301	248,240	198,770	179,059	247,164
Predicted Deaths upto T_{end}(P)	5,852 +/- 704	533 +/- 60	4,104 +/- 479	236 +/- 27	38,513 +/- 4,312	26,698 +/- 2,968	9,069 +/- 1,043	30,348 +/- 3,546	33,343 +/- 3,705
Actual Deaths upto T_{end}(A)	6,066	515	5,732	192	41,384	28,340	9,062	30,180	35,134
Predicted Cases/Million population upto T_{end}(P)	2602 +/- 313	1832 +/- 206	2235 +/- 260	1453 +/- 163	2752 +/- 308	5175 +/- 575	2185 +/- 251	2265 +/- 265	3943 +/- 438
Actual Cases/Million population upto T_{end}(A)	2,828	1,774	7,375	1,367	4,806	5,293	2,395	2,673	4,092
Predicted CFR (deaths/million) upto T_{end}(P)	338 +/- 41	92 +/- 10	402 +/- 47	44 +/- 5	578 +/- 65	569 +/- 63	109 +/- 13	453 +/- 53	552 +/- 61
Actual CFR (deaths/million) upto T_{end}(A)	351	89	562	36	621	604	109	450	582

Supplementary Figures

Fig. S1a

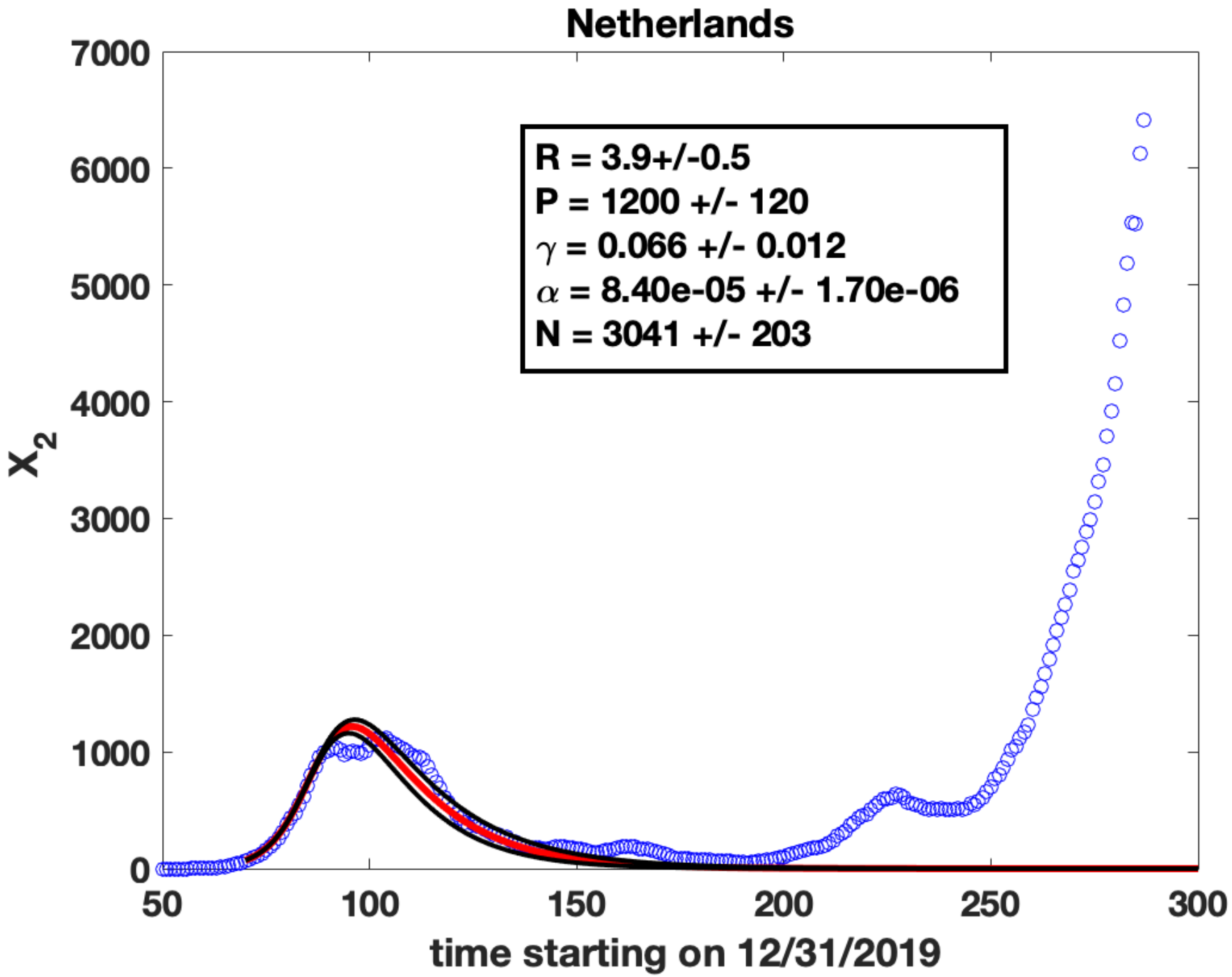

Netherlands
R = 3.9+/-0.5
P = 1200 +/- 120
γ = 0.066 +/- 0.012
α = 8.40e-05 +/- 1.70e-06
N = 3041 +/- 203
X2
7000
6000
5000
4000
3000
2000
1000
0
50
100
150
200
250
300
time starting on 12/31/2019

Fig. S1b

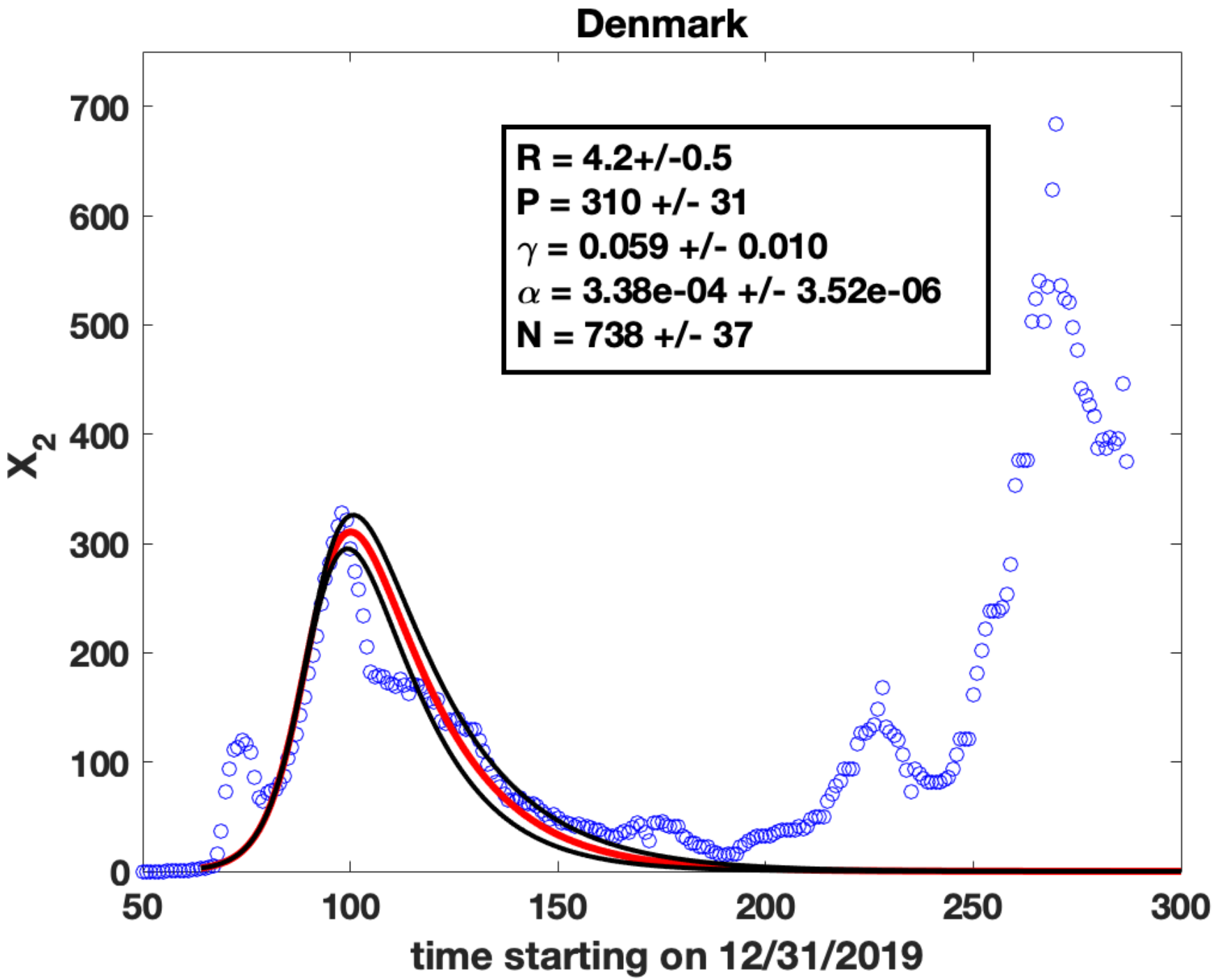

Denmark
R = 4.2+/-0.5
P = 310 +/- 31
γ = 0.059 +/- 0.010
α = 3.38e-04 +/- 3.52e-06
N = 738 +/- 37
X2
0
100
200
300
400
500
600
700
50
100
150
200
250
300
time starting on 12/31/2019

Fig. S1c

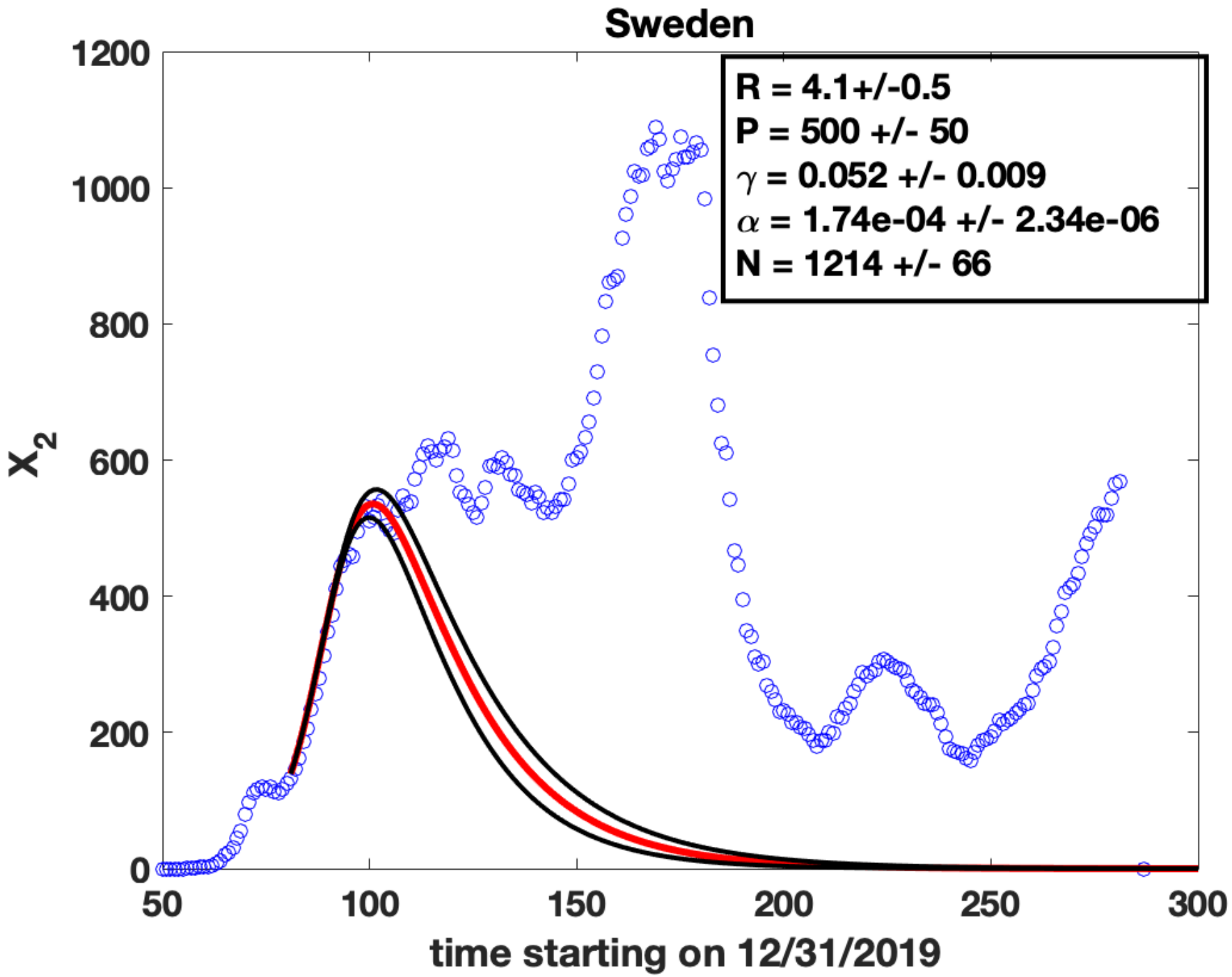

Sweden
R = 4.1+/-0.5
P = 500 +/- 50
γ = 0.052 +/- 0.009
α = 1.74e-04 +/- 2.34e-06
N = 1214 +/- 66
X2
1200
1000
800
600
400
200
0
50
100
150
200
250
300
time starting on 12/31/2019

Fig. S1d

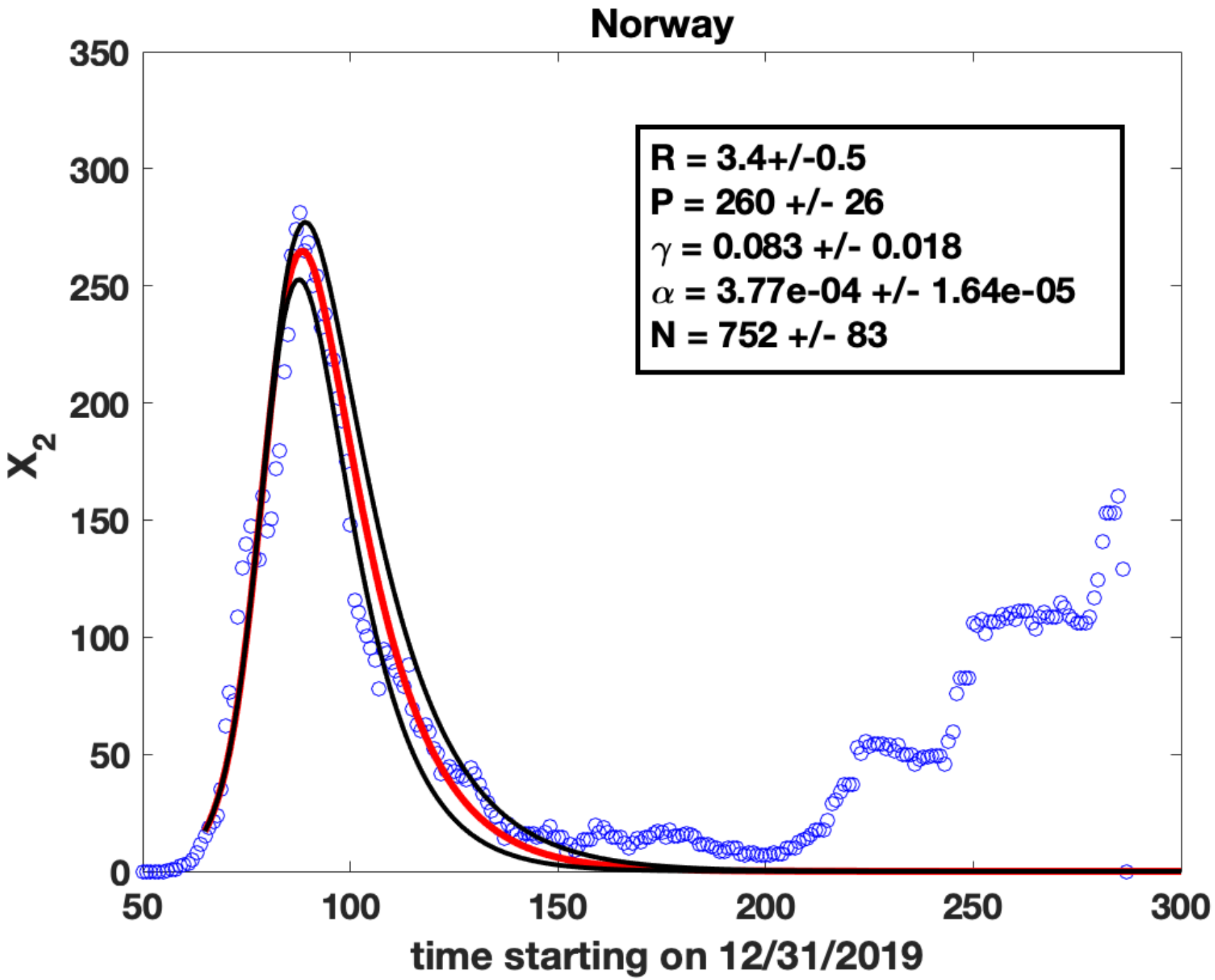

Norway
R = 3.4+/-0.5
P = 260 +/- 26
γ = 0.083 +/- 0.018
α = 3.77e-04 +/- 1.64e-05
N = 752 +/- 83
X_2
time starting on 12/31/2019
0
50
100
150
200
250
300
350
50
100
150
200
250
300

Fig. S1e

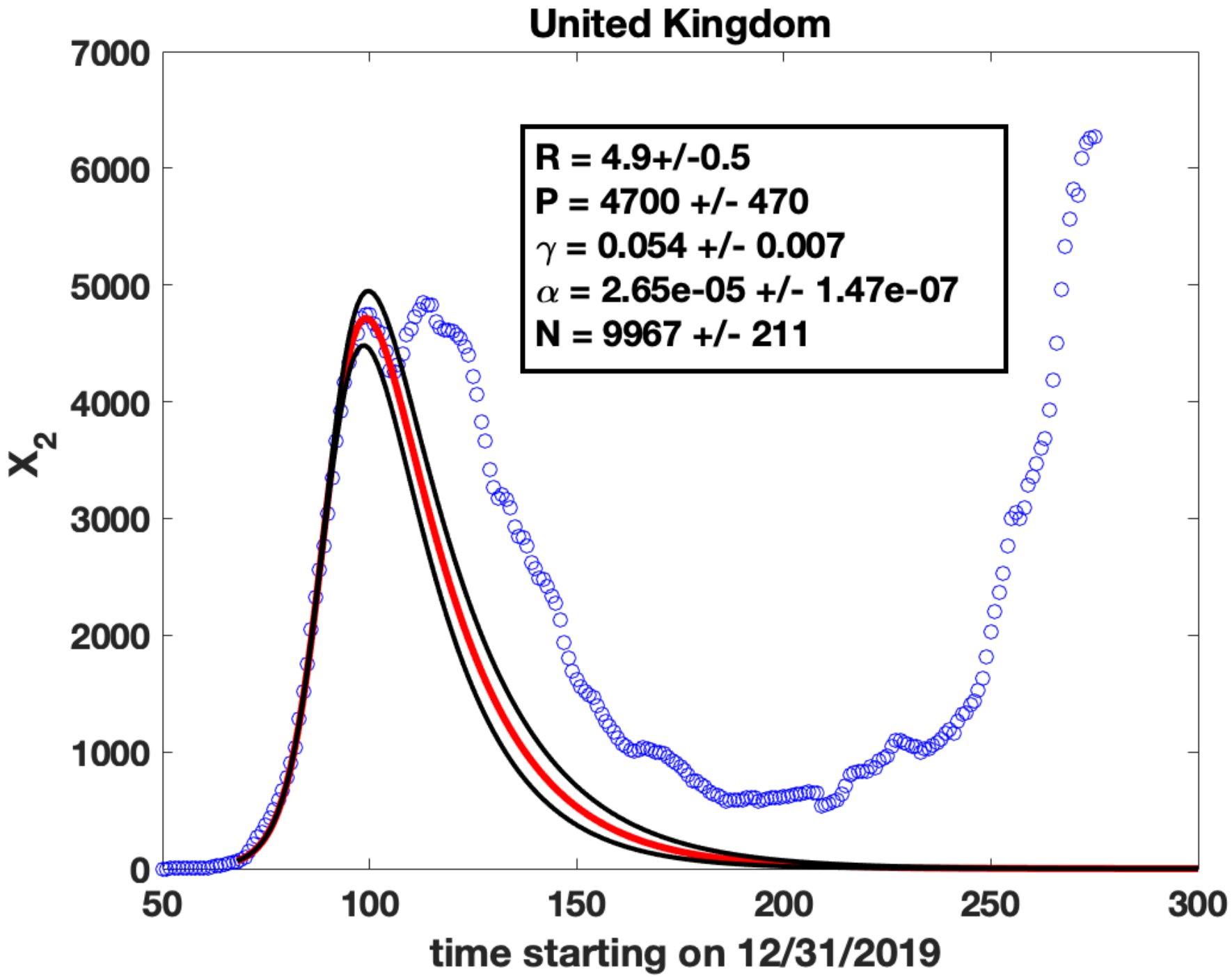

United Kingdom
R = 4.9+/-0.5
P = 4700 +/- 470
γ = 0.054 +/- 0.007
α = 2.65e-05 +/- 1.47e-07
N = 9967 +/- 211
X2
7000
6000
5000
4000
3000
2000
1000
0
50
100
150
200
250
300
time starting on 12/31/2019

Fig. S1f

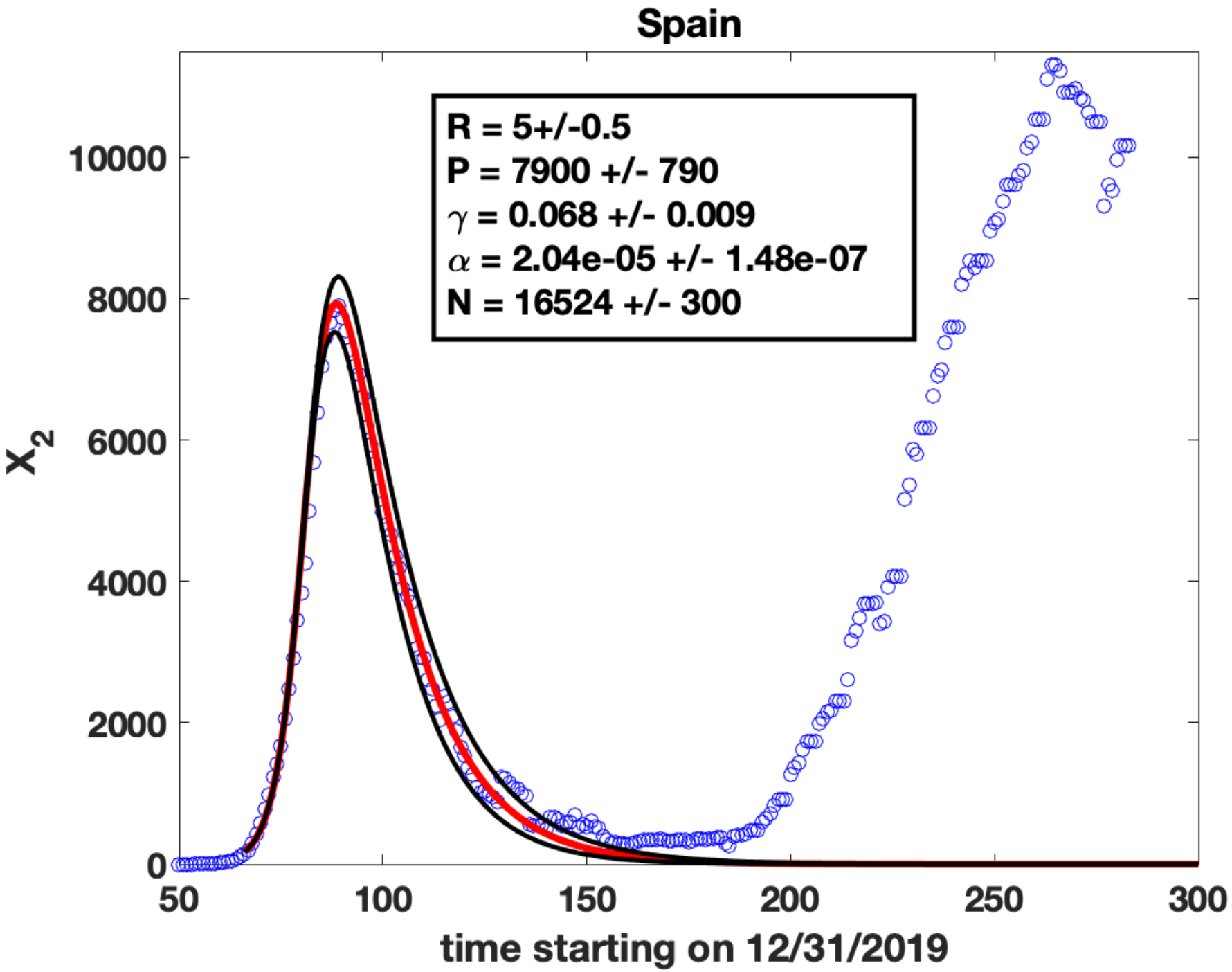

Spain
R = 5+/-0.5
P = 7900 +/- 790
γ = 0.068 +/- 0.009
α = 2.04e-05 +/- 1.48e-07
N = 16524 +/- 300
10000
8000
6000
4000
2000
0
X2
50
100
150
200
250
300
time starting on 12/31/2019

Fig. S1g

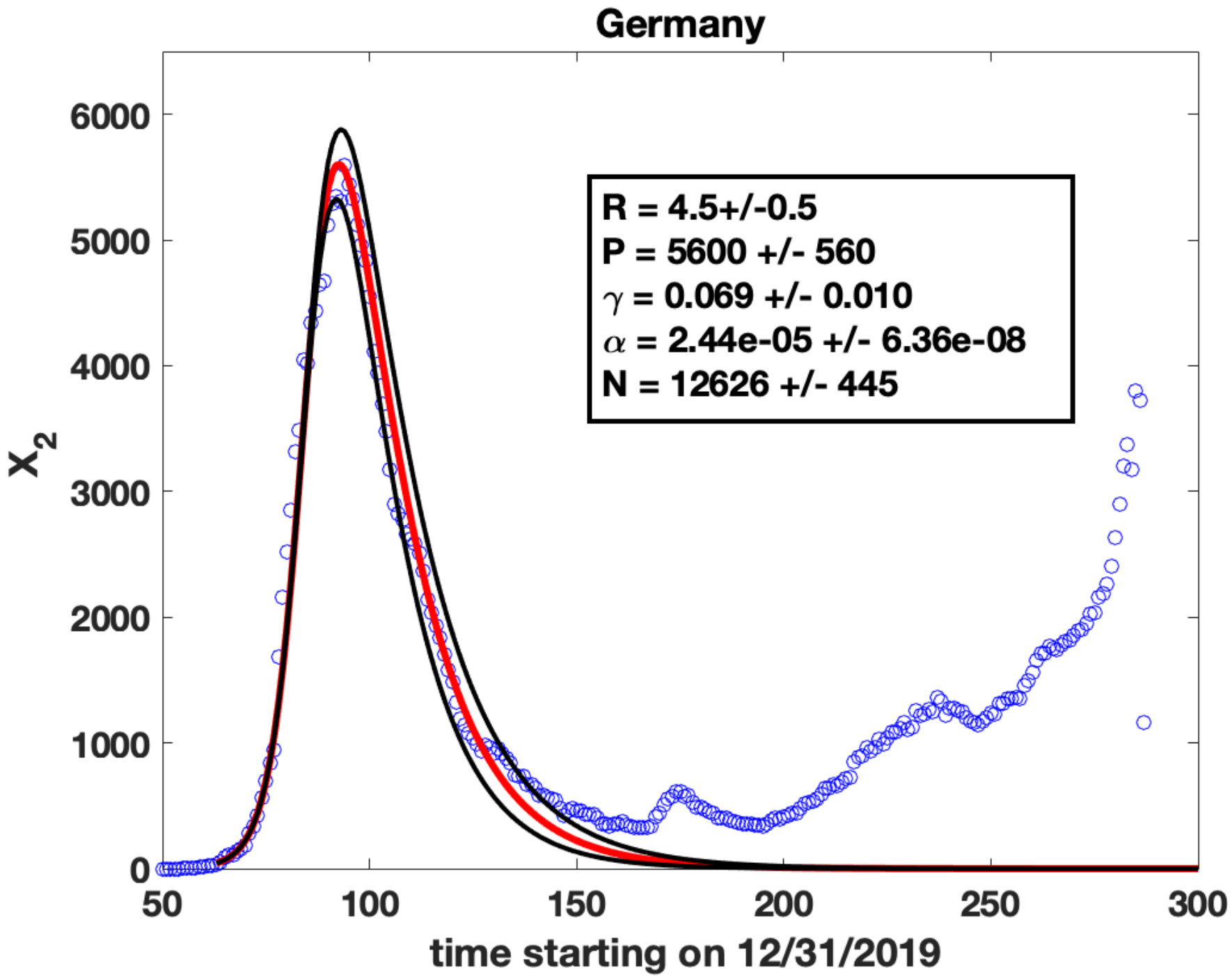

Germany
R = 4.5+/-0.5
P = 5600 +/- 560
γ = 0.069 +/- 0.010
α = 2.44e-05 +/- 6.36e-08
N = 12626 +/- 445
X2
6000
5000
4000
3000
2000
1000
0
50
100
150
200
250
300
time starting on 12/31/2019

Fig. S1h

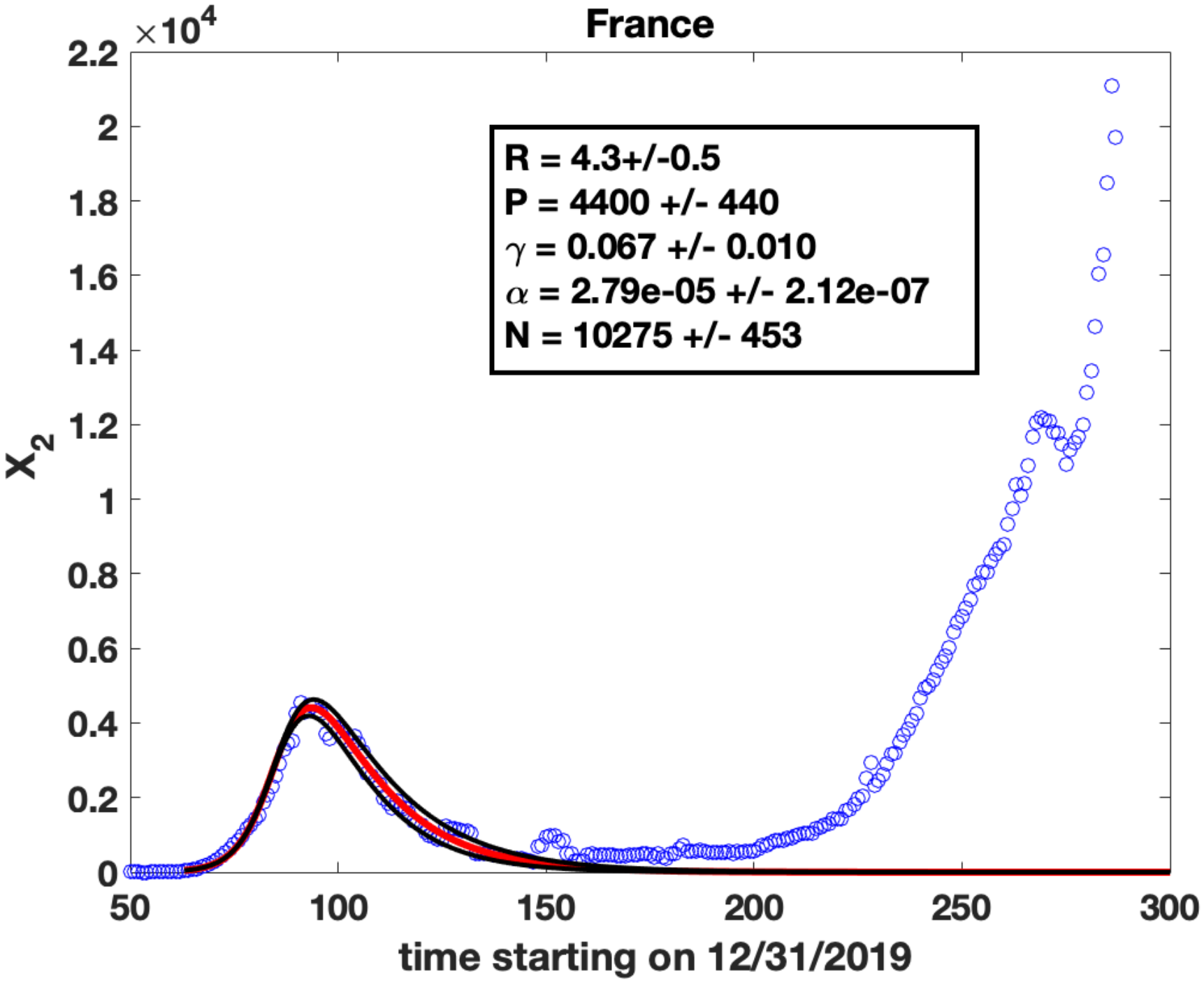
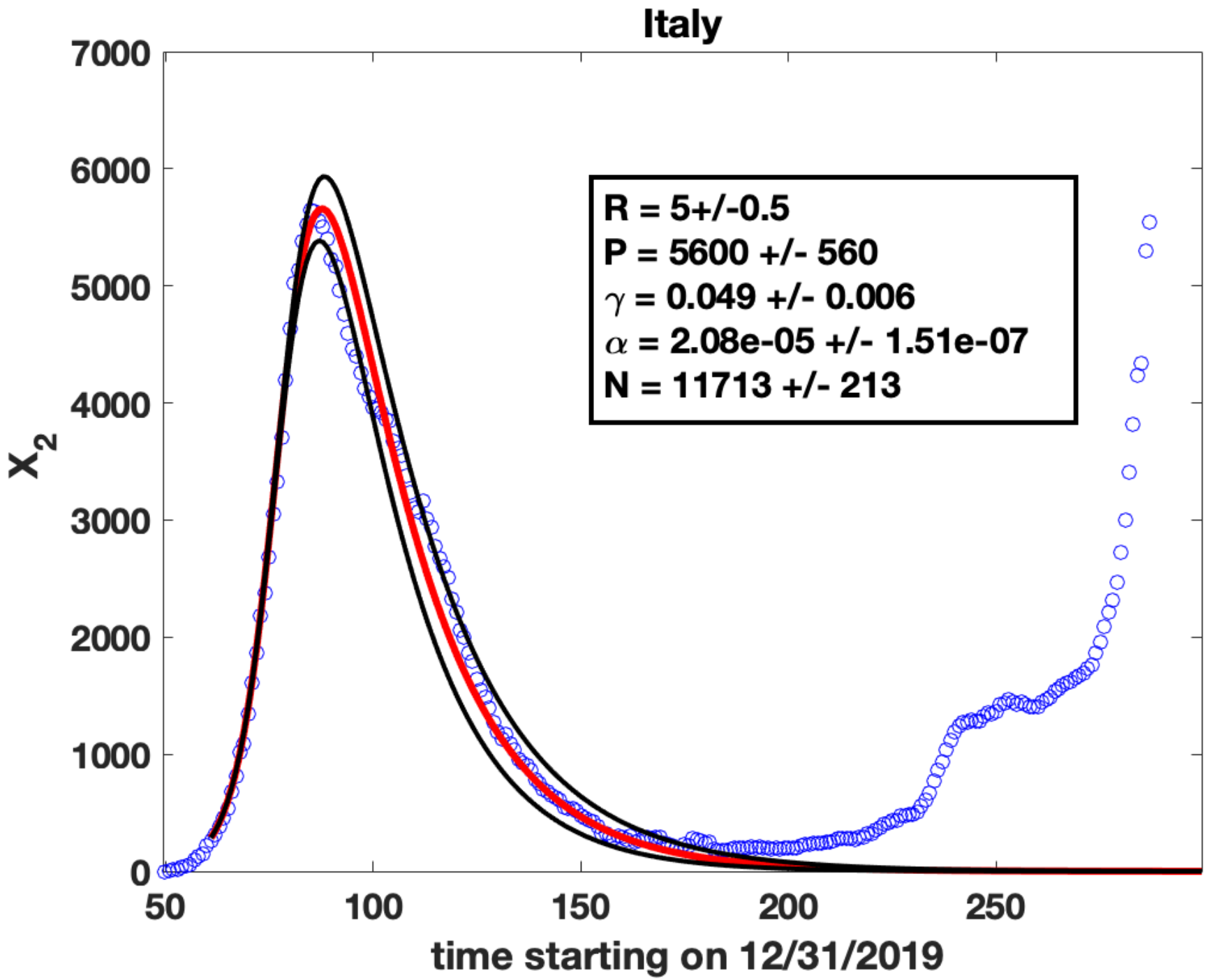

France
×10^4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
X_2
R = 4.3+/-0.5
P = 4400 +/- 440
γ = 0.067 +/- 0.010
α = 2.79e-05 +/- 2.12e-07
N = 10275 +/- 453
50
100
150
200
250
300
time starting on 12/31/2019

Fig. S1i


Italy
R = 5+/-0.5
P = 5600 +/- 560
γ = 0.049 +/- 0.006
α = 2.08e-05 +/- 1.51e-07
N = 11713 +/- 213
X2
7000
6000
5000
4000
3000
2000
1000
0
50
100
150
200
250
time starting on 12/31/2019