I&EC research
Industrial & Engineering Chemistry Research

pubs.acs.org/IECR

Article

Use of a Modified SIRD Model to Analyze COVID-19 Data

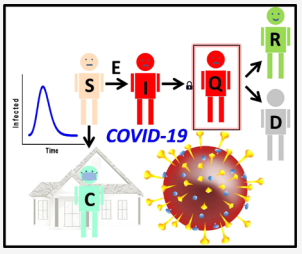
Devosmita Sen and Debasis Sen*

Cite This: *Ind. Eng. Chem. Res.* 2021, 60, 4251−4260

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Since the starting of the year 2020, the whole world is facing a challenge due to an outbreak of an unprecedented COVID-19 pandemic owing to a novel coronavirus. Here, a modified susceptible−infected−recovered−dead model has been used to analyze the time series data of the pandemic for five countries. It is established that the present model is capable of simultaneously explaining the temporal evolution of active-infected, recovered, and dead population of all these five countries. The key parameters governing the temporal evolution of the spread of this pandemic are estimated and compared.



1. INTRODUCTION

Since the beginning of 2020, the whole world has been experiencing a major and unprecedented global crisis, owing to the outbreak of the COVID-19 pandemic.[1−3] The infection is resulting in severe, and sometimes even fatal, respiratory diseases such as acute respiratory distress syndrome.[4] Such an infectious disease, with a humongous social and economic impact was never seen before, at least in the recent past.[5] COVID-19 arose due to a strain of a novel coronavirus that has rapidly spread throughout the globe,[1,6] originating from and infecting a large number of people in Wuhan, China.[7−9] The spread of this disease has a complex time dependence, which is governed not by the number of infected people alone, but is strongly correlated with aspects such as total population of the country, various norms and measures taken by the nation at a particular time, and many more.[10−14] Because of a lack of previous experience in controlling a similar pandemic with such a high impact in the recent past, it is difficult to anticipate the size of the population that may get affected by this pandemic and the typical time required for its control.

The abovementioned crisis immediately calls for a quantitative understanding of the time evolution of this complex and non-linear process through computer modeling. Statistical and mathematical analysis of reported data can provide valuable insights into the trend of the spread and thus can assist in planning various social measures to contain the spread of the virus as quickly as possible. Further, analysis of reported epidemic data plays a vital role in analyzing the underlying phenomena involved in spreading of the disease and to make predictions about future trends. This enables various organizations to efficiently plan their steps toward containing this spread.

In literature, a few models have been proposed to explain such data. These can primarily be classified into two categories, collective models[12,13,15−18] and networked models.[19−23] Some examples of the former class of models are growth and logistic models,[12] the susceptible−infected−recovered−dead (SIRD) model, and their modifications,[9,16] collectively termed as compartmental models. At this juncture, it is worthy to mention that such a phenomenon has a close resemblance to kinetics of chemical reactions in general,[24] where transition from one state to another is associated with a specific rate. In epidemic modeling, the corresponding rates may be expressed in terms of the instantaneous number of infected, recovered people, and so forth. Owing to the complex interdependence of several processes governing the spread of infections and recovery, a proper mathematical model should be able to simultaneously predict the temporal behavior of infected, recovered, and dead people. As the protection procedures demand quarantine, confinement, social distancing lockdown measures, and so forth,[26] a simple SIRD model only gives a preliminary understanding of the process and is not sufficient to describe such complex processes in general. Here, we model the spread of the present pandemic using a generalized SIRD model, taking into account the fraction of population which is exposed, under quarantine, confined, active-infected, recovered, and expired at time instant *t*. It is noteworthy that proper modeling calls for a simultaneous corroboration of the time evolution of all the three independent reported data sets, namely, active number of infections and cumulative number of recovered and expired people due to the infection.

In this manuscript, we report the analysis of COVID-19 time series data for five countries—China, Italy, France, the United States, and India (and two of its highly affected states)—using the modified SIRD model. We have shown that this model is capable of explaining the current data significantly well for all

Received: September 25, 2020
Revised: December 23, 2020
Accepted: January 20, 2021
Published: February 2, 2021

https://dx.doi.org/10.1021/acs.iecr.0c04754
Ind. Eng. Chem. Res. 2021, 60, 4251−4260

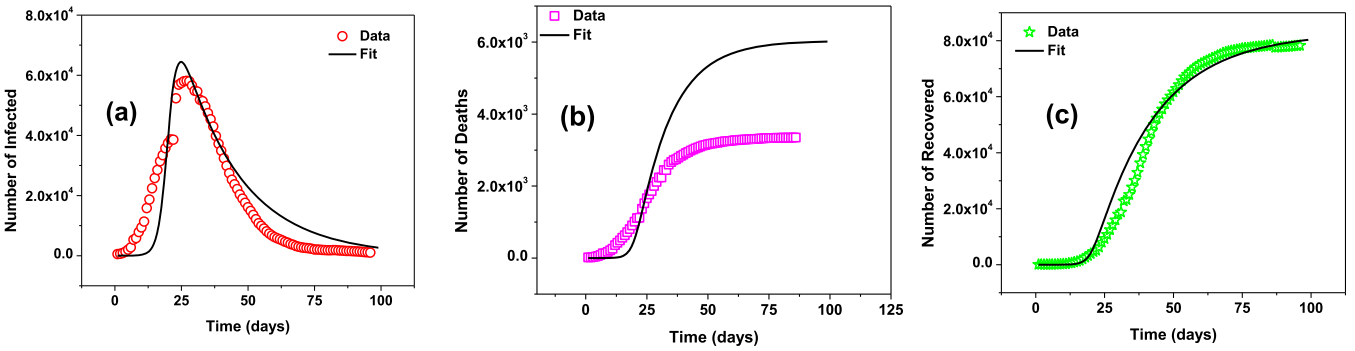


Figure 1. (a) Number of active cases of infection in China vs time in days. (b) Cumulative number of deaths due to the infection vs time in days. (c) Cumulative number of people recovered from the infection vs time in days. Solid lines represent the model (SIRD) fit to the data.

these five countries. It should be noted that the present approach is unique because it is capable of explaining simultaneously all the three reported data sets: active, recovered, and dead population. The data have been obtained from https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases which is compiled by the Johns Hopkins University Center for Systems Science and Engineering (JHU CCSE) from various sources including the World Health Organization (WHO) and from reported data on Indian state websites https://arogya.maharashtra.gov.in/1175/Novel--Corona-Virus, https://gujcovid19.gujarat.gov.in/.

2. MODEL

2.1. Standard SIRD Model. The SIRD model deals with only the number of susceptible, infected, recovered, and dead people. According to this model, a susceptible person in contact with an infected person is prone to get infected.[25,27,28] An infected person can either recover from the disease or die due to the infection. Thus, this model considers that the sum of $S(t)$, $I(t)$, $R(t)$, and $D(t)$ remains constant. Further, it is assumed that all people who are exposed to the virus get infected immediately, that is, there is no latent time between the exposure and infection. No effect of confinement or quarantine is considered.

The governing equations of this model are

$$I + S \xrightarrow{\beta} I \qquad \frac{dS(t)}{dt} = -\beta\frac{I(t)S(t)}{N} \tag{1}$$

$$I \xrightarrow{\lambda} R \qquad \frac{dI(t)}{dt} = \beta\frac{I(t)S(t)}{N} - \lambda I(t) - k_d I(t) \tag{2}$$

$$I \xrightarrow{k_d} D \qquad \frac{dR(t)}{dt} = \lambda I(t) \tag{3}$$

$$\frac{dD(t)}{dt} = k_d I(t) \tag{4}$$

$S(t)$—susceptible population, people who can be infected by the virus at a particular time t. $I(t)$—people who have been infected by the virus at a particular time t, that is, active number of cases of infection. $R(t)$—cumulative number of people at time t who have recovered from the infection. $D(t)$—cumulative number of people at time t who have died due to the infection.

In Figure 1, we present the data from China and the fit of this model with the initial number of infected, recovered, and dead people as 650, 0, and 1, respectively. As evident from the graphs below, although this model gives a trend of the features exhibited by the time-series data, it hardly represents the actual trends.

As the effect of latent time has not been considered, growth in active cases of infections, as predicted by the SIRD model, remains very steep. Further, as quarantine effects have not been considered, the decay predicted here is much slower than reality. The predicted value of total number of deaths is also much higher than actual.

Hence, this model needs proper modifications to corroborate all the three data sets—infected, recovered, and dead—simultaneously. Containment, lockdown, and quarantine measures and asymptomatic population play a significant role in the way the infection spreads over time. Hence, we need to incorporate these factors into the model.

2.2. Modified SIRD Model. In this model, we consider the effect of exposure, quarantine, confinement, and asymptomatic population, which are the cases in reality.[25,26,29] Almost all countries around the globe have adopted strategies to control[30] the spread of this virus. Some of these include lockdown measures, wearing face masks, quarantining infected people, and so forth. There is a latent period after which the symptoms of the disease manifest. Further, there exist some asymptomatic cases too. Hence, to incorporate all these effects, we modify the SIRD model as below.

The governing equations of the present model are

$$S \xrightarrow{\alpha} C \qquad \frac{dS(t)}{dt} = -\alpha S(t) - \beta\frac{I(t)S(t)}{N} - \sigma\frac{S(t)A(t)}{N} - \eta S(t) \tag{5}$$

$$I + S \xrightarrow{\beta} E \tag{6}$$

$$C \xrightarrow{\mu} E \qquad \frac{dA(t)}{dt} = -\tau A(t) + \xi E(t) \tag{7}$$

$$S \xrightarrow{\eta} E \quad S + A \xrightarrow{\sigma} E \qquad \frac{dC(t)}{dt} = \alpha S(t) - \mu C(t) \tag{8}$$

$$E \xrightarrow{\xi} A \quad E \xrightarrow{\gamma} I \qquad \frac{dE(t)}{dt} = -\gamma E(t) + \beta\frac{I(t)S(t)}{N} + \mu C(t) + \eta S(t) + \sigma\frac{S(t)A(t)}{N} - \xi E(t) \tag{9}$$

$$A \xrightarrow{\tau} I \quad I \xrightarrow{\delta} Q \qquad \frac{dI(t)}{dt} = \tau A(t) + \gamma E(t) - \delta I(t) \tag{10}$$

https://dx.doi.org/10.1021/acs.iecr.0c04754
Ind. Eng. Chem. Res. 2021, 60, 4251−4260

$$Q \xrightarrow{\lambda(t)} R \qquad \frac{dR(t)}{dt} = \lambda(t)Q(t) \tag{11}$$

$$Q \xrightarrow{k_d(t)} D$$

$$\frac{dQ(t)}{dt} = \delta I(t) - \lambda(t)Q(t) - k_d(t)Q(t) \tag{12}$$

$$\frac{dD(t)}{dt} = k_d(t)Q(t) \tag{13}$$

$$S(t) + C(t) + E(t) + A(t) + I(t) + Q(t) + R(t) + D(t) = N \tag{14}$$

where N is the total population. $C(t)$ represents the confined population, who maintain social distancing norms, wear preventive face masks, and follow lockdown rules. $E(t)$ represents the exposed population, those people who have been exposed to the virus but have not been tested positive for infection yet. The exposed population $E(t)$ is considered to be non-infectious, that is, they do not transmit the infection to others at time t. This term arises to account for a latent time between exposure to the virus and the infection. $A(t)$ represents the asymptomatic population, that is, those people who have been exposed to the virus but do not show any explicit symptom of infection. They have not yet been tested positive for the virus and thus are not part of the infected compartment. However, the asymptomatic population $A(t)$ is considered to transmit the infection to others. This term thus contributes to an increase in the number of infection. $Q(t)$ is the quarantined population at time t. It is considered that people who are infected are eventually quarantined and do not come in contact with susceptible people. Hence, quarantined people, in principle, do not contribute to the spread of the infection any more. There is a delay between a person getting infected and getting quarantined, which is known as the quarantine time. α (day^{-1}) is the protection rate which quantifies the rate at which the susceptible population becomes confined. β (day^{-1}) is known as the exposure rate and represents the average rate of spread of infection due to contact of a susceptible person with an infected person. μ (day^{-1}) is the rate constant for conversion of the confined population to the exposed compartment probably by air-borne spread or through surface infection. η (day^{-1}) is the rate constant for conversion of the susceptible population directly to an exposed population compartment (without contact with the infected person), again probably through air-borne and surface infection mechanisms. σ (day^{-1}) is the rate constant for conversion of the susceptible population to the exposed compartment through contact with asymptomatic persons. ξ (day^{-1}) is the rate constant for exposed to asymptomatic conversion. τ (day^{-1}) is the rate constant for the asymptomatic population moving into the infected compartment. γ (day^{-1}) represents the average rate at which the exposed population becomes infected. γ^{-1} is the average latent time, that is, the average time for an already exposed person to get infected. δ (day^{-1}) is the quarantine rate which is the rate at which the infected population is quarantined. δ^{-1} is the average quarantine time which is equal to the average time delay between infection and quarantine. $\lambda(t)$ (day^{-1}) represents the recovery rate. This is the rate at which the infected population recovers from infection. Different measures control this parameter and thus λ is often time-dependent. $k_d(t)$ (day^{-1}) represents the death rate due to infection/disease. This rate can also be time-dependent based on several factors such as acquired immunity, treatment, and so forth.

In the present case, the functional form of λ and k_d have been considered as $\lambda(t) = \lambda_0 \exp(-\lambda_1 t)$ and $k_d(t) = k_{d0} \exp(-k_{d1} t)$, respectively.

The scheme of the present model is presented in Figure 2. According to this model, the whole population is susceptible to the infection at time $t = 0$. As more and more people follow lockdown and social distancing or other preventive measures, they move to the confined population (C) with a rate of α.

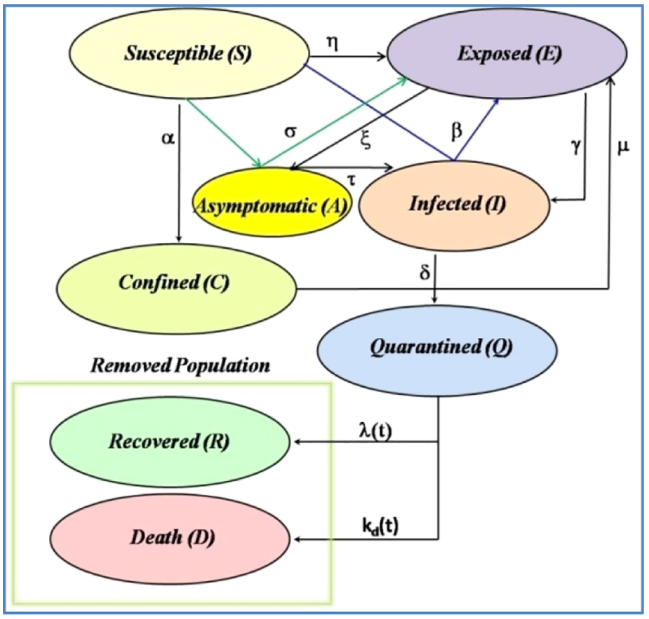


Figure 2. Schematic representation of the present model.

In the next step, a susceptible person in contact with an infected person can get exposed to the virus with a rate constant of β and contact with an asymptomatic person with a rate constant of σ. The confined people may get affected either by air-borne or surface spread, which demands for a rate at which the confined population moves to the exposed population (μ). Further, a susceptible person can get infected either by air-borne or surface infection with a rate constant η. There is a time delay between the viral exposure and infection, and the average time taken for this is denoted by γ^{-1}. An exposed person moves to the asymptomatic compartment with a rate constant of ξ. As more and more people are tested for viral infection, those who are asymptomatic move to the infected compartment as they are now identified as infected. This happens with a rate constant of τ. The infected population is eventually quarantined with a rate δ. A quarantined person can either recover or expire due to infection. The rates associated with these are given by λ and k_d, respectively. It is assumed that a person recovering from the infection acquires sufficient immunity and is not re-infected soon. The natural birth or death rate is much lesser than the death rate due to this disease and thus has been neglected while modeling the spread of infection.

The solutions of the set of the differential equations and their fitting to the data demand estimating the unknown parameters, which are α, β, μ, η, σ, ξ, τ, γ, δ, λ_0, λ_1, k_{d0}, and k_{d1} in this case. Further, the initial conditions for the variables corresponding to each compartment are also similarly crucial. The initial values are known for at least $S(t)$, $I(t)$, $R(t)$, and $D(t)$ from the reported

https://dx.doi.org/10.1021/acs.iecr.0c04754
Ind. Eng. Chem. Res. 2021, 60, 4251−4260

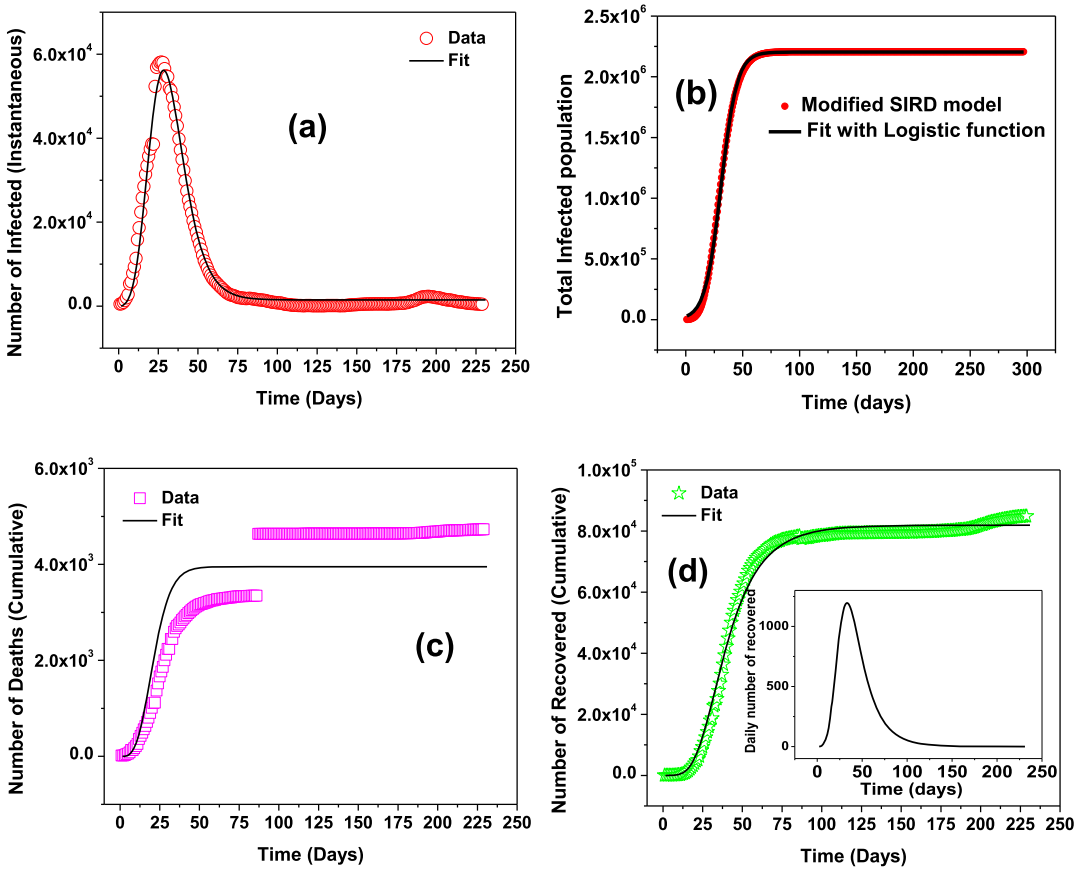


Figure 3. (a) Number of active cases of infection in China vs time in days. (b) Fit of the standard logistic function with the simulated cumulative cases of infection. (c) Cumulative number of deaths due to infection vs time in days. (d) Cumulative number of people who recovered from the infection vs time in days. Solid lines represent the model (modified SIRD) fit to the data. The inset shows the daily number of recovered persons.

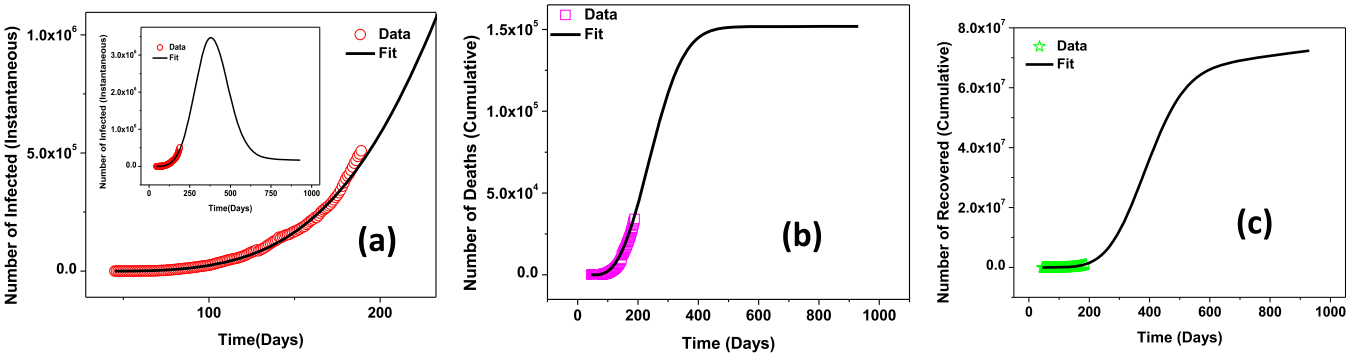


Figure 4. (a) Number of active cases of infection in India vs time in days. The inset shows the trend of infection as predicted by the model. (b) Cumulative number of deaths due to infection vs time in days. (c) Cumulative number of people who recovered from the infection vs time in days. Solid lines represent the model (modified SIRD) fit to the data.

data. However, choosing the initial conditions for $C(t)$, $A(t)$, $E(t)$, and $Q(t)$ is also important and tricky as the results are somewhat sensitive to these values. It was observed during the fitting that α, β, γ, and δ are more crucial than the other rate constants as far as the sensitivities are concerned.

3. PARAMETER ESTIMATION

To avoid complexity, often a set of differential equations are solved by approximating them into different equations and solving them algebraically.[31] However, owing to the statistical fluctuations in data, this method may result in some uncertainties in some cases. A more realistic approach is probably to solve the differential equations and fit the model with the data by a suitable method.

In the present case, the above set of differential equations has been solved numerically and the model was fitted to the data of infected, recovered, and death cases simultaneously using a non-linear least square. A Python program was written for this purpose and the following χ^2 was minimized during the fitting process

https://dx.doi.org/10.1021/acs.iecr.0c04754
Ind. Eng. Chem. Res. 2021, 60, 4251−4260

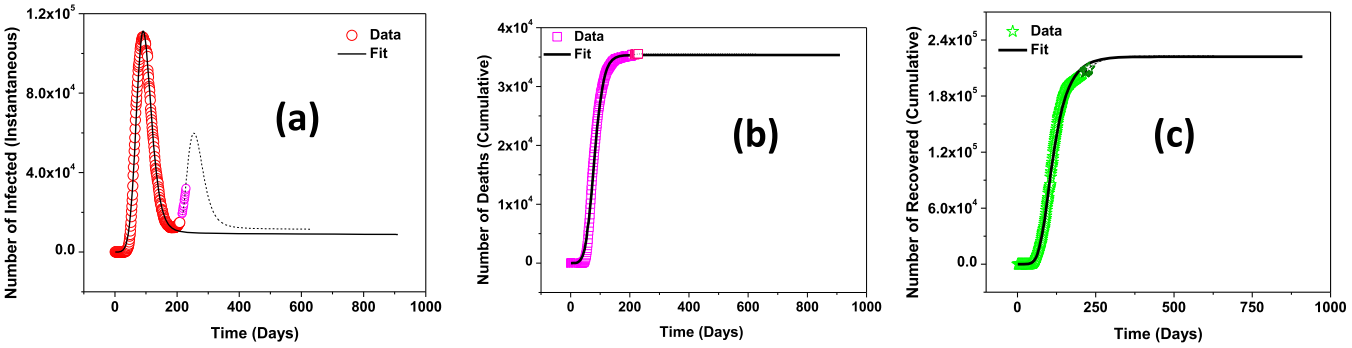


Figure 5. (a) Number of active cases of infection in Italy vs time in days. (b) Cumulative number of deaths due to infection vs time in days. (c) Cumulative number of people who recovered from the infection vs time in days. Solid lines represent the model (modified SIRD) fit to the data.

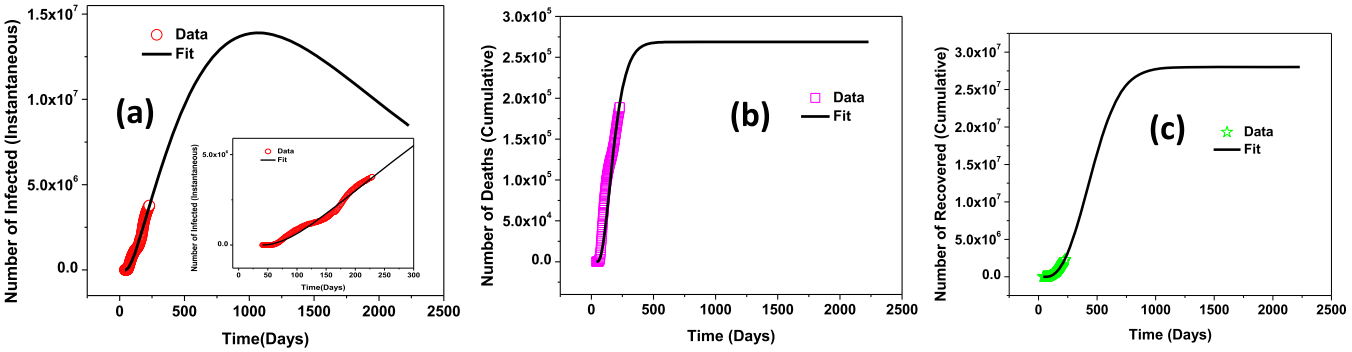


Figure 6. (a) Number of active cases of infection in the United States vs time. (b) Cumulative number of deaths due to infection vs time. (c) Cumulative number of people who recovered from the infection vs time. Solid lines represent the model (modified SIRD) fit to the data.

$$\chi^2\ (\text{parameters}) = \left(\left[\sum_t (I_{\text{data}}(t) - I_{\text{model}}(t))^2 + (R_{\text{data}}(t) - R_{\text{model}}(t))^2 + (D_{\text{data}}(t) - D_{\text{model}}(t))^2\right]/(n-m)\right)^{1/2} \quad (15)$$

where n represents the number of data points and m is the number of parameters to be solved for.

4. RESULTS AND DISCUSSION

In Figures 3−6, we depict the COVID-19 time series data sets from China, India, Italy, and the United States, respectively, along with the corresponding simultaneous fit of this model for the infected, recovered, and dead population. It is seen from all these figures that this model is indeed capable of representing the temporal behavior of all the three data sets nicely for all the countries. Here, it is to be noted that the chosen data sets are such that for some countries, the number of active cases has crossed its peak value while for rest of the countries, it is yet to do so. It gives a confidence on the validity of the model fitting. Below, we will discuss about their trends.

From Figure 3a,c,d (data of China), it is evident that the spread of infection is well past its peak after 25−30 days, and daily new cases of infection have reduced significantly. The number of deaths has reached saturation level. It is interesting to note that death surges suddenly at around 85 days but stays nearly constant after that. It is observed that it takes a total of ∼100 days for the infections to get reduced significantly. From the reported data of the dead population in China, it is evident that there was a sudden step-like jump after ∼90 days. As the present model takes into account only continuous variations in the external factors and thus in the temporal behavior of the spread, the model represents only a qualitative fit in Figure 3c. However, it is interesting to observe that even for such a sudden jump in the data, the model effectively mimics its principal trend. In literature, the total number of infected people has been modeled using the logistic growth model.[12,32] This is particularly because when the number of infected persons remains small, the disease spreads slowly. Further, a similar situation happens if the susceptible population is small. The disease grows fast only when both the number of infected people and the number of susceptible people are simultaneously significant, as this enables contact between the two types of populations. It is worthy to mention that for a standard exponential growth, the rate is proportional to the instantaneous number of infected people. However, as the infection spreads widely, the infected population becomes comparable to the susceptible population. As the mechanism of spread is through the contact of an infected person with susceptible population, the growth rate of infection becomes proportional to the product of infected and susceptible population. This gives the expression for the logistic growth model (eq 16). We have compared our simulated results for cumulative number of infected people with the logistic function as given below.

$$y(t) = \frac{a}{1 + e^{-k(t-t_c)}} \quad (16)$$

For China, the values of a, k, and t_c are 2.2×10^6, 0.1428 (days^{-1}), and 31 (days) respectively. The parameter a represents the total number of people infected by this pandemic. t_c corresponds to the time from the initiation of infection at which a peak in active number of cases is observed. For other countries considered in this work, we have also compared the trend of time variation for the total infected population as

https://dx.doi.org/10.1021/acs.iecr.0c04754
Ind. Eng. Chem. Res. 2021, 60, 4251−4260

Table 1. Estimated Parameters from the Fit of the Modified SIRD Model with Data

	country						
	China	India			Italy	the United States	France
		overall	Maharashtra	Gujarat			
population ($\times 10^7$)	143.93	138.73	12.2	6.5	6.046	33.1	6.527
population density (per km^2)	153	382	365	308	206	36	119
α (day^{-1})	0.080	0.0025	0.0046	0.00149	0.0194	0.01215	0.003709
β (day^{-1})	0.912	0.180	0.2237	0.1132	7.567	7.6678	6.09694
μ (day^{-1})	6.675×10^{-8}	8.152×10^{-6}	9.284×10^{-6}	1.092×10^{-5}	2.278×10^{-6}	3.514×10^{-3}	2.3655×10^{-13}
η (day^{-1})	1.068×10^{-7}	1.529×10^{-8}	7.689×10^{-8}	2.207×10^{-6}	9.180×10^{-7}	1.489×10^{-2}	9.2679×10^{-17}
σ (day^{-1})	0.166	9.217×10^{-6}	1.00×10^{-10}	2.510×10^{-7}	1.463×10^{-3}	1.00×10^{-15}	1.00×10^{-17}
τ (day^{-1})	4.802	4.221×10^{-7}	1.061×10^{-7}	4.215×10^{-6}	1.109×10^{-4}	9.1234×10^{-4}	1.0×10^{-17}
ξ (day^{-1})	1.097	2.200×10^{-10}	5.3046×10^{-9}	0.0153	0.263	1.535×10^{-3}	2.2366×10^{-2}
γ (day^{-1})	0.720	0.023	0.0665	0.0629	0.021	9.8076×10^{-5}	0.02282
δ (day^{-1})	0.105	0.076	0.067	0.0711	0.077	9.7942×10^{-3}	0.0639
λ_0 (day^{-1})	0.216	0.6079	0.308	8.8107	0.157	3.1848	4.0796
λ_1 (day^{-1})	0.055	1.00×10^{-8}	3.487×10^{-8}	0.0086	0.025	1.1072×10^{-2}	0.0597
k_{d0} (day^{-1})	0.035	0.0267	0.1255	1.9142	0.779	3.52629	4.2616
k_{d1} (day^{-1})	0.078	0.0166	0.0167	0.02419	0.061	2.718×10^{-2}	0.06927
initial infected	1	1	1	1	1	1	1
initial exposed	1	1	1	1	1	1	1
initial confined	0	0	0	0	0	0	0
initial quarantined	1	1	1	1	1	1	1

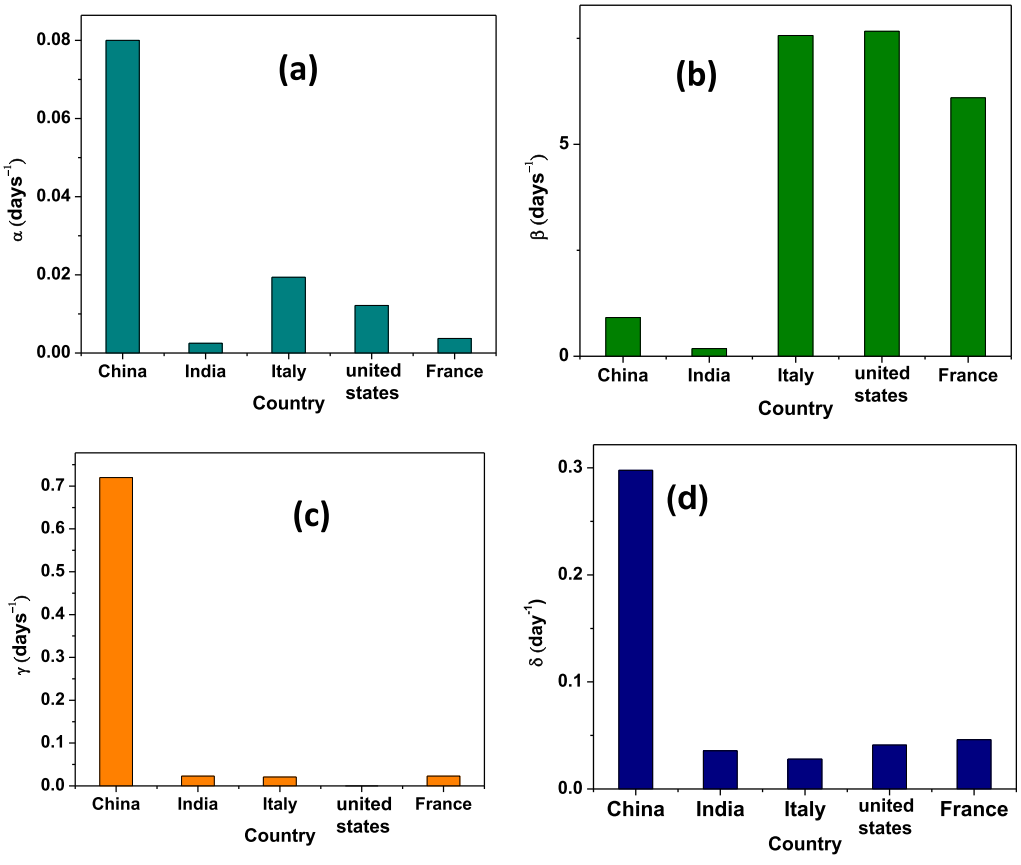


Figure 7. Comparison of estimated parameters α, β, γ, and δ for different countries.

obtained from the present model to the logistic growth model and the results are shown in the Supporting Information (Figure S4). It is evident from the figures that in reality, the complexity of the problem often does not follow such a standard logistic model. It is seen that the behavior in a few countries/states deviates significantly from logistic growth unlike that in China. This shows that our present modified SIRD model is capable of explaining the data even when the growth does not strictly follow a logistic behavior.

For India, the infection is still increasing and the model suggests that infections should reach its peak ~300 days after the initiation of infection and it will take a longer time (more than

https://dx.doi.org/10.1021/acs.iecr.0c04754
Ind. Eng. Chem. Res. 2021, 60, 4251−4260

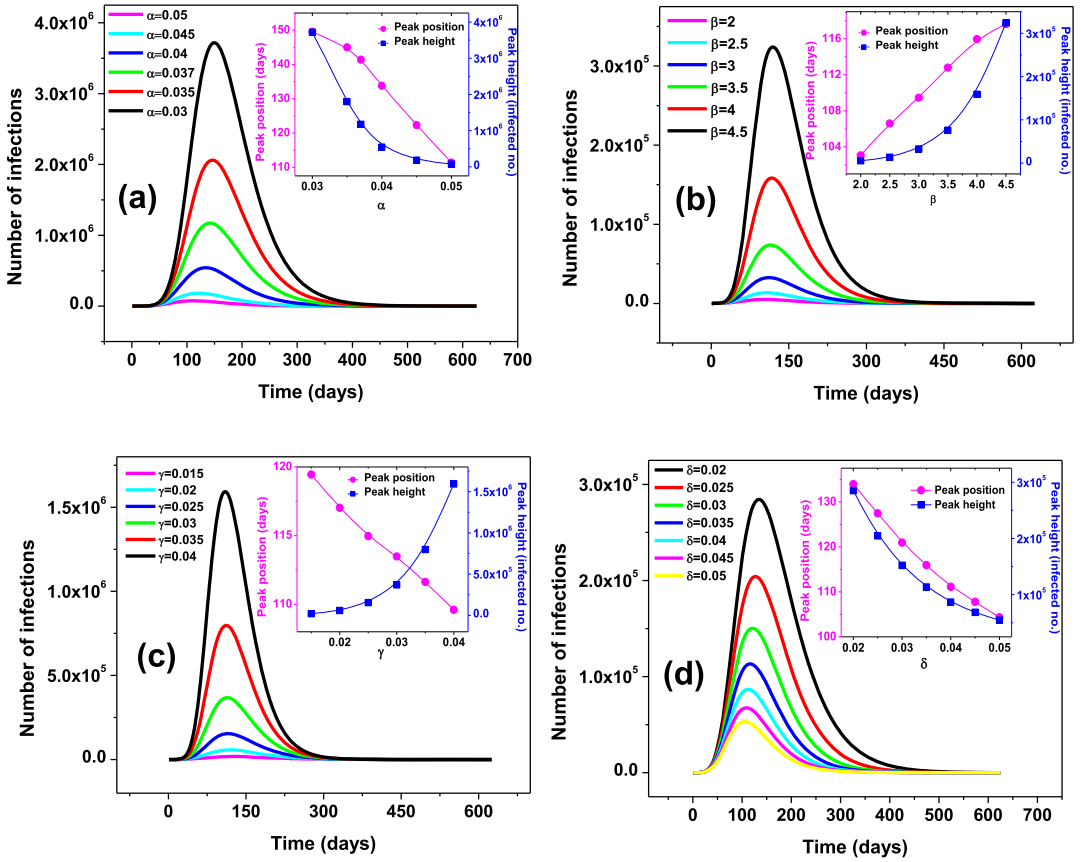


Figure 8. Simulated curve shows the trend of sensitivity of the parameters (a) α, (b) β, (c) γ, and (d) δ, showing the sensitivity of these parameters. The inset in each graph represents the variation of peak height and peak position with the respective parameter.

700 days) for infections to die down significantly as opposed to the situation in China. It is anticipated that more than 5×10^8 people will ultimately get infected. The number of deaths is still increasing and should start saturating after ~500 days. The total number of deaths is expected to be close to 1.75×10^5.

Within India, two states, namely, Maharashtra and Gujarat, have been severely affected. The model fit for Maharashtra state predicts that the peak of the active number of infections is expected to be at ~250 days from initiation of infection and is expected to continue for a relatively longer time (almost 600 days). This is depicted in Figure S1 (Supporting Information). Data from Gujarat state (Figure S2) show that the infection is approaching its peak value after nearly 300 days and it will take more than 600 days for the infection to significantly die down.

Data from Italy show that it too has gone past its first peak (~60 days), but the number of active cases of infection has started increasing again and is approaching a second peak after 200 days. It is expected to take more than 300 days for the number of active cases of the second wave of infection to reduce significantly from the start of second infection. This shows that the latent time (γ^{-1}) for infection in Italy is greater than that in China. This is also evident by comparing the values of γ obtained while fitting the model for Italy and China. This is represented in Table 1 and Figure 7.

The scenario of infection in the United States is somewhat complex in nature than the cases in China and Italy. Here, the functionality of growth rate of the active cases appears to be quite non-uniform with some oscillatory behavior although the infection is still growing on an average. The overall peak is not achieved yet. The present model considers an average picture of such oscillatory data and suggests that the infection should continue even beyond 500 days. It is discernible that a broad peak for active infection is expected only after 1000 days. The number of deaths is still increasing and there is a signature that it would saturate slowly only beyond 500 days. The model predicts that a total of more than 10^7 people will get infected by this pandemic. This analysis shows that the United States has the highest infection rate, followed by Italy and then France (Table 1).

Data from France (Figure S3) indicate that the number of active cases of infection has been at its first peak at ~100 days from its start. However, France also is approaching the next peak of infection, which is expected beyond 300 days. It is observed that the growth of infections has been very steep. This is also evident by comparing the value of β obtained from the fitting, as represented in Table 1 and Figure 7. The saturation level of number of deaths is expected to be around 3.5×10^4.

In Figure 7, the time-independent parameters α, β, γ, and δ have been compared as bar charts for these countries. It is observed that γ is highest in China, followed by France, and lowest in the United States. This indicates that the average latent time (γ^{-1}) between exposure and infection is significantly less for China. This is evident by the fact that the curve for growth of infections is quite steep. Quarantine rate (δ) is also significantly higher in China than in all the other countries considered here. This is probably because, as the complicated infections suddenly originated in China, the severity of these infections were immediately anticipated and quarantine measures were taken. A high value of δ for China indicates a small quarantine time. China has the highest value of α and β is the highest in the

https://dx.doi.org/10.1021/acs.iecr.0c04754
Ind. Eng. Chem. Res. 2021, 60, 4251−4260

United States. Here, it should be noted that the actual cause behind the variation of the parameters for these countries is complex enough to differentiate as the physicochemical parameters, life styles, natural/acquired immunity, temperature, food habit, and preventative social measures highly vary among these countries. Of course, it opens up a new arena for research in general.

5. SENSITIVITY ANALYSIS

α represents the rate at which the susceptible population moves into the confined population by following lockdown and social distancing. If α increases, the fraction of population that is susceptible decreases and thus reduces the chances of spread of infection. Hence, as α increases, the number of infections decreases. The sensitivity of the parameter α is depicted in Figure 8a. Further, as α increases, the peak of infection shifts leftward. This can be explained by the fact that if the rate of confinement is more, then less number of people will be infected, and the infection dies down faster, and so, the peak of active number of cases of infections is attained early. Hence, if most people are in the compartment of the confined population, infection can be controlled quickly. This explains the necessity of implementing lockdown measures and maintaining social distancing norms.

As the exposure rate β increases, there is an increase in the number of infected people. Due to this, the spread of infection would also increase, and infection is expected to sustain longer. Hence, the curve for active number of cases of infection continues to grow, and it takes a lot of time for infection to reach its maximum value. In such a situation, the peak of the curve shifts rightward, as given by Figure 8b.

As γ increases, number of infections increases. This is because, as the latent time (γ^{-1}) decreases, exposed people get infected more quickly, and this contributes to an increase in the number of active infections at a particular time. Further, from Figure 8c it is seen that as γ^{-1} decreases, vis-à-vis as γ increases, the position of peak of infection shifts toward the left. This is because even though in the initial period of growth of infections, the active number of cases remains high, the rate of spread of infections (β) remains unchanged. Hence, new cases of infections decrease and most infected people recover. Hence, infections die out very soon and when γ increases, the peak of infections is attained early.

As the quarantine rate δ increases, the number of infections decreases, as given by Figure 8d. This is expected because quarantining infected people implies that they no more infect susceptible people and thus the number of infections can be significantly reduced. The peak of the curve shifts leftward because of the same reason as in the case of α. As more and more infected people get quarantined, the spread of disease can be controlled faster and the infection dies down quickly. Hence, as δ increases, the peak of the curve shifts left. As evident from Figure 8d, there is a significant shift in the peak of the curve of active number of infections, and this explains the crucial role of quarantine for controlling such a pandemic.

In brief, an increase in α (confinement rate) leads to a decrease in the number of infections. Its peak is attained early and the infection dies down faster. An increase in β (exposure rate) leads to an increase in the number of infections and it takes a longer time to attain a maximum. An increase in γ (inverse of latent time) leads to an increase in the active number of infections too, but it dies out faster. An increase in δ (quarantine rate) leads to a decrease in the number of infections, and the peak is attained earlier.

The variation of quarantine rate with latent time is depicted in Figure 9. It shows that if latent time (γ^{-1}) is short, then

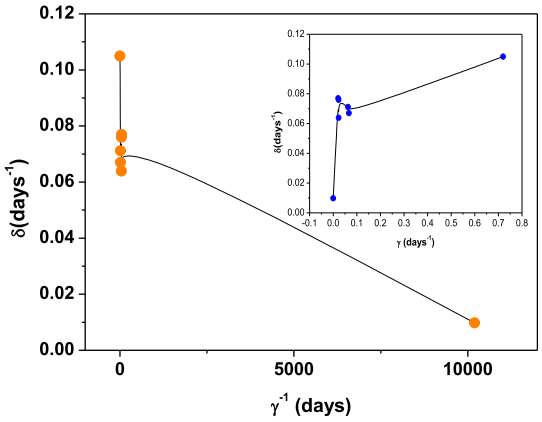


Figure 9. Quarantine rate (δ) vs latent time (γ^{-1}). The inset shows the variation of quarantine time (δ^{-1}) with latent time (γ^{-1}).

quarantine rate (δ) is more. This is because if symptoms appear within a very short time after exposure, then people will be more aware and thus diagnosed effectively and quarantined faster. However, if latent time is high, then quarantine is also likely to be delayed and hence quarantine rate will be low. Physically, this can be explained by the fact that if symptoms of infection are visible only a long time after a person is exposed, then the person will not be aware of it for a long time, and this may lead to delay in testing even after symptoms start showing up. This would lead to further delay in identifying and quarantining the infected people, leading to an increase in the time between infection and quarantine. This shows that there is an inverse relationship between quarantine rate and latent time.

In Figure 10, the variation of the significant parameters with the population density of the respective country/state is

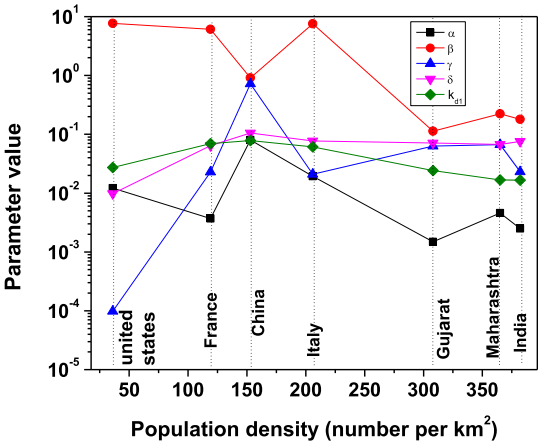


Figure 10. Variation of the parameters with the population density of the countries/states.

depicted. It is evident that, except for China, the parameters show a fair variation trend with population density. γ shows an incremental trend with density while β shows a somewhat decremental trend. It is interesting to note that for China, the values of γ and α show a maximum while β shows a local

https://dx.doi.org/10.1021/acs.iecr.0c04754
Ind. Eng. Chem. Res. 2021, 60, 4251−4260

minimum. This corroborates with the fastest containment of spread in this particular country.

At this juncture, it is worth mentioning a few possible limitations of this model with regard to the pandemic data. From the above analysis, it was realized that there can be an oscillatory behavior in the reported data (such as in Figures 6 or S2), which probably arises due to variation in the confinement and lockdown conditions. This effect has to be explicitly imposed in the model and cannot be directly integrated into the equations. In such a situation, the model needs to be solved in parts. Further, the model cannot handle a situation with an abrupt surge in death/infection/recovery data. However, in such a situation, the model is indeed well capable of explaining an average picture of the evolution. It is to be noted that owing to the complexity of the governing equations, it was indispensable to consider several parameters in the present model, which were estimated through a non-linear least square fitting. In this work, the fitting was implemented in several iterations, where the value of a parameter obtained from the previous iteration was considered as the guess value for the next iteration until the χ^2 value reached a minimum. Moreover, only the four crucial parameters (α, β, γ, and δ) were kept free till the last iteration. The uncertainty on the estimated parameters was noted. In all the cases, it was ensured that this uncertainty value remained insignificant and much less than 1%.

6. CONCLUSIONS

- standard SIRD model has been modified to analyze the temporal evolution of spread of infection because of the present COVID-19 pandemic for five countries across the globe.
- It is found that this model is capable of explaining the spread of all the three reported data sets, active, recovered, and dead, simultaneously for these countries.
- This model quantifies the role of confinement and quarantine in controlling the spread and also takes into account the practical consideration of asymptomatic populations, surface/air-borne infection possibilities, and so forth.
- It is shown from the analysis of the data from five countries that a decrease in the latent time leads to an increase in the quarantine rate, thus maintaining an inverse relationship.
- Infection rate (β), latent time (γ^{-1}), and quarantine rate (δ) are found to vary significantly from one country to the other.
- This work indicates that while the model is equally applicable in explaining the pandemic in various countries, the variation in the rate constants is significant for different countries. This is possibly because of inherent differences in life styles, natural/acquired immunity, environmental conditions, food habits, preventative social measures, and so forth, and this needs special attention for future research.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.iecr.0c04754.

Data and fit for Maharashtra and Gujarat states of India, France, and comparison of time dependence of the total infected population with the logistic model (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Debasis Sen − *Solid State Physics Division, Bhabha Atomic Research Centre, Mumbai 400085, India; Homi Bhabha National Institute, Mumbai 400094, India;* orcid.org/0000-0002-9080-0866; Email: debasis@barc.gov.in

Author

Devosmita Sen − *Department of Chemical Engineering, Indian Institute of Technology, Bombay, Mumbai 400076, India*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.iecr.0c04754

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) WHO. *W.H.O. Coronavirus Disease 2019 (COVID-19). Situation Report*, 2020; Vol. *127*, pp 1−17.

(2) Zhou, P.; Yang, X.-L.; Wang, X.-G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.-R.; Zhu, Y.; Li, B.; Huang, C.-L.; Chen, H.-D.; Chen, J.; Luo, Y.; Guo, H.; Jiang, R.-D.; Liu, M.-Q.; Chen, Y.; Shen, X.-R.; Wang, X.; Zheng, X.-S.; Zhao, K.; Chen, Q.-J.; Deng, F.; Liu, L.-L.; Yan, B.; Zhan, F.-X.; Wang, Y.-Y.; Xiao, G.-F.; Shi, Z.-L. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *579*, 270−273.

(3) Nicholas, C. P. Viral Infections and the Development of Disinfection: 100 Years of Progress at I&ECR. *Ind. Eng. Chem. Res.* **2020**, *59*, 6345−6346.

(4) Chen, N.; Zhou, M.; Dong, X.; Qu, J.; Gong, F.; Han, Y.; Qiu, Y.; Wang, J.; Liu, Y.; Wei, Y.; Xia, J. A.; Yu, T.; Zhang, X.; Zhang, L. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* **2020**, *395*, 507−513.

(5) Nicola, M.; Alsafi, Z.; Sohrabi, C.; Kerwan, A.; Al-Jabir, A.; Iosifidis, C.; Agha, M.; Agha, R. The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *Int. J. Surg.* **2020**, *78*, 185−193.

(6) Yuan, J.; Li, M.; Lv, G.; Lu, Z. K. Monitoring transmissibility and mortality of COVID-19 in Europe. *Int. J. Infect. Dis.* **2020**, *95*, 311−315.

(7) Wu, Z.; McGoogan, J. M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72,314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* **2020**, *323*, 1239−1242.

(8) Velavan, T. P.; Meyer, C. G. The COVID-19 epidemic. *Trop. Med. Int. Health* **2020**, *25*, 278−280.

(9) Giordano, G.; Blanchini, F.; Bruno, R.; Colaneri, P.; Di Filippo, A.; Di Matteo, A.; Colaneri, M. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nat. Med.* **2020**, *26*, 855−860.

(10) Remuzzi, A.; Remuzzi, G. COVID-19 and Italy: what next? *Lancet* **2020**, *395*, 1225−1228.

(11) Zhao, S.; Lin, Q.; Ran, J.; Musa, S. S.; Yang, G.; Wang, W.; Lou, Y.; Gao, D.; Yang, L.; He, D.; Wang, M. H. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. *Int. J. Infect. Dis.* **2020**, *92*, 214−217.

(12) Chowell, G. Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts. *Infect. Dis. Model.* **2017**, *2*, 379−398.

(13) Kermack, W. O.; McKendrick, A. G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. London, Ser. A* **1927**, *115*, 700−721.

(14) Calafiore, G.; Novara, C.; Possieri, C. A Modified SIR Model for the COVID-19 Contagion in Italy. **2020**, arXiv:2003.14391 [physics.soc-ph].

https://dx.doi.org/10.1021/acs.iecr.0c04754
Ind. Eng. Chem. Res. 2021, 60, 4251−4260

(15) Chowell, G.; Tariq, A.; Hyman, J. M. A novel sub-epidemic modeling framework for short-term forecasting epidemic waves. *BMC Med.* **2019**, *17*, 164.
(16) Yang, Z.; Zeng, Z.; Wang, K.; Wong, S.-S.; Liang, W.; Zanin, M.; Liu, P.; Cao, X.; Gao, Z.; Mai, Z.; Liang, J.; Liu, X.; Li, S.; Li, Y.; Ye, F.; Guan, W.; Yang, Y.; Li, F.; Luo, S.; Xie, Y.; Liu, B.; Wang, Z.; Zhang, S.; Wang, Y.; Zhong, N.; He, J. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *J. Thorac. Dis.* **2020**, *12*, 165−174.
(17) Li, L.; Yang, Z.; Dang, Z.; Meng, C.; Huang, J.; Meng, H.; Wang, D.; Chen, G.; Zhang, J.; Peng, H.; Shao, Y. Propagation analysis and prediction of the COVID-19. *Infect. Dis. Model.* **2020**, *5*, 282−292.
(18) Lin, Q.; Zhao, S.; Gao, D.; Lou, Y.; Yang, S.; Musa, S. S.; Wang, M. H.; Cai, Y.; Wang, W.; Yang, L.; He, D. A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action. *Int. J. Infect. Dis.* **2020**, *93*, 211−216.
(19) Keeling, M. J.; Eames, K. T. D. Networks and epidemic models. *J. R. Soc., Interface* **2005**, *2*, 295−307.
(20) Nadini, M.; Rizzo, A.; Porfiri, M. Epidemic spreading in temporal and adaptive networks with static backbone. *IEEE Trans. Network Sci. Eng.* **2020**, *7*, 549−561.
(21) Nowzari, C.; Preciado, V. M.; Pappas, G. J. Optimal resource allocation for control of networked epidemic models. *IEEE Trans. Control Network Syst.* **2015**, *4*, 159−169.
(22) Piontti, A. Y. P.; Gomes, M. F. D. C.; Samay, N.; Perra, N.; Vespignani, A. The infection tree of global epidemics. *Network Sci.* **2014**, *2*, 132−137.
(23) Pastor-Satorras, R.; Castellano, C.; Van Mieghem, P.; Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **2015**, *87*, 925.
(24) Fogler, H. S. *Elements of Chemical Reaction Engineering*, 5th ed.; Prentice Hall International Series in the Physical and Chemical Engineering Sciences; Pearson Education (US): United States, 2016; Chapter 9, pp 333−384.
(25) Wang, P.; Jia, J. Stationary distribution of a stochastic SIRD epidemic model of Ebola with double saturated incidence rates and vaccination. *Adv. Differ. Equ.* **2019**, *2019*, 433.
(26) Binti Hamzah, F. A.; Hau, C.; Nazri, H.; Ligot, D.; Lee, G.; Shaib, M.; Zaidon, U.; Abdullah, A.; Chung, M.; Ong, C.; Chew, P. CoronaTracker: World-wide COVID-19 Outbreak Data Analysis and Prediction. *Bull. W.H.O.* **2020**, *1*, 32.
(27) Anastassopoulou, C.; Russo, L.; Tsakris, A.; Siettos, C. Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS One* **2020**, *15*, e0230405.
(28) Yuan, M.; Yin, W.; Tao, Z.; Tan, W.; Hu, Y. Association of radiologic findings with mortality of patients infected with 2019 novel coronavirus in Wuhan, China. *PLoS One* **2020**, *15*, e0230548.
(29) Tsay, C.; Lejarza, F.; Stadtherr, M. A.; Baldea, M. Modeling, state estimation, and optimal control for the US COVID-19 outbreak. *Sci. Rep.* **2020**, *10*, 10711.
(30) Prem, K.; Liu, Y.; Russell, T. W.; Kucharski, A. J.; Eggo, R. M.; Davies, N.; Jit, M.; Klepac, P.; Flasche, S.; Clifford, S.; Pearson, C. A. B.; Munday, J. D.; Abbott, S.; Gibbs, H.; Rosello, A.; Quilty, B. J.; Jombart, T.; Sun, F.; Diamond, C.; Gimma, A.; van Zandvoort, K.; Funk, S.; Jarvis, C. I.; Edmunds, W. J.; Bosse, N. I.; Hellewell, J. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *Lancet Public Health* **2020**, *5*, e261−e270.
(31) Cantó, B.; Coll, C.; Sánchez, E. Estimation of parameters in a structured SIR model. *Adv. Differ. Equ.* **2017**, *2017*, 33.
(32) Pell, B.; Kuang, Y.; Viboud, C.; Chowell, G. Using phenomenological models for forecasting the 2015 Ebola challenge. *Epidemics* **2018**, *22*, 62−70.

https://dx.doi.org/10.1021/acs.iecr.0c04754
Ind. Eng. Chem. Res. 2021, 60, 4251−4260