Modeling the COVID-19 Pandemic Using an SEIHR Model With Human Migration

NIU, Ruiwu; WONG, Eric W. M.; CHAN, Yin-Chi; VAN WYK, Michael Antonie; CHEN, Guanrong

Published in:
IEEE Access

Published: 01/01/2020

Document Version:
Final Published version, also known as Publisher’s PDF, Publisher’s Final version or Version of Record

License:
CC BY-NC-ND

Publication record in CityU Scholars:
Go to record

Published version (DOI):
10.1109/ACCESS.2020.3032584

Publication details:

Citing this paper
Please note that where the full-text provided on CityU Scholars is the Post-print version (also known as Accepted Author Manuscript, Peer-reviewed or Author Final version), it may differ from the Final Published version. When citing, ensure that you check and use the publisher’s definitive version for pagination and other details.

General rights
Copyright for the publications made accessible via the CityU Scholars portal is retained by the author(s) and/or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights. Users may not further distribute the material or use it for any profit-making activity or commercial gain.

Publisher permission
Permission for previously published items are in accordance with publisher's copyright policies sourced from the SHERPA RoMEO database. Links to full text versions (either Published or Post-print) are only available if corresponding publishers allow open access.

Take down policy
Contact lbscholars@cityu.edu.hk if you believe that this document breaches copyright and provide us with details. We will remove access to the work immediately and investigate your claim.
Modeling the COVID-19 Pandemic Using an SEIHR Model With Human Migration

RUIWU NIU¹, ERIC W. M. WONG⁰¹, (Senior Member, IEEE), YIN-CHI CHAN⁰¹, (Member, IEEE), MICHAËL ANTONIE VAN WYK⁰², (Senior Member, IEEE), AND GUANRONG CHEN⁰¹, (Life Fellow, IEEE)

¹Department of Electrical Engineering, City University of Hong Kong, Hong Kong, China
²School of Electrical and Information Engineering, University of the Witwatersrand at Johannesburg, Johannesburg 2000, South Africa

Corresponding author: Eric W. M. Wong (eeewong@cityu.edu.hk)

This work was supported in part by the Health and Medical Research Fund, in part by the Food and Health Bureau, and in part by The Government of the Hong Kong Special Administrative Region, China, under Grant 16171921.

ABSTRACT The 2019 novel coronavirus disease (COVID-19) outbreak has become a worldwide problem. Due to globalization and the proliferation of international travel, many countries are now facing local epidemics. The existence of asymptomatic and pre-symptomatic transmissions makes it more difficult to control disease transmission by isolating infectious individuals. To accurately describe and represent the spread of COVID-19, we suggest a susceptible-exposed-infected-hospitalized-removed (SEIHR) model with human migrations, where the “exposed” (asymptomatic) individuals are contagious. From this model, we derive the basic reproduction number of the disease and its relationship with the model parameters. We find that, for highly contagious diseases like COVID-19, when the adjacent region’s epidemic is not severe, a large migration rate can reduce the speed of local epidemic spreading at the price of infecting the neighboring regions. In addition, since “infected” (symptomatic) patients are isolated almost immediately, the transmission rate of the epidemic is more sensitive to that of the “exposed” (asymptomatic) individuals. Furthermore, we investigate the impact of various interventions, e.g. isolation and border control, on the speed of disease propagation and the resultant demand on medical facilities, and find that a strict intervention measure can be more effective than closing the borders. Finally, we use some real historical data of COVID-19 caseloads from different regions, including Hong Kong, to validate the modified SEIHR model, and make an accurate prediction for the third wave of the outbreak in Hong Kong.

INDEX TERMS COVID-19, modified SEIHR model, disease transmission model, disease control, human migration.

I. INTRODUCTION

The Coronavirus Disease 2019 (COVID-19) pandemic has resulted in over 34.4 million reported cases and 1.02 million deaths throughout 188 countries and territories (as of 3 October 2020) [1] and has caused great concern among governments, the World Health Organization, and scientists worldwide over the past few months. The outbreak of COVID-19 has been more rapid and widespread than the Severe Acute Respiratory Syndrome (SARS) outbreak in 2003 and the Middle East Respiratory Syndrome (MERS) outbreaks in 2012 in Saudi Arabia [2], [3] and 2015 in South Korea. If stringent intervention measures are not taken to restrain the pandemic, COVID-19 may eventually reach the same level of devastation as the “Spanish Flu”, which infected 500 million people and caused 17–50 million deaths worldwide [4].

One reason why COVID-19 spreads so rapidly is that infectious individuals are contagious in the latent period, and a significant proportion of infected individuals do not show any symptoms throughout the entire course of the disease [5]. Since these cases are extremely difficult to detect and isolate, they can easily cause what are known as pre-symptomatic and asymptomatic transmissions, respectively, making it much harder to control the outbreak. Furthermore, the exponential increase of the numbers of patients in most regions has had a devastating impact on healthcare systems worldwide, further increasing the already high death rates. Therefore, it is
extremely important to monitor the spreading processes of COVID-19 and study its medical and social impacts.

A. MATHEMATICAL MODELS FOR EPIDEMIOLOGY

In the early 20th century, Ross [6] established a mathematical model for the transmission of malaria between humans and mosquitoes, and proposed the concept of a threshold value for the spreading of the disease. This concept was further refined by MacDonald [7] who proposed what has become known as the basic reproduction number $R_0$ of a disease. $R_0$ has the property that an epidemic will persist if $R_0 > 1$ but will diminish if $R_0 < 1$. Therefore, the derivation of $R_0$ is a key step in the development of various epidemiological models.

A definitive work on epidemiological models is due to Kermack and McKendrick [8], [9], who proposed compartmental models to describe the number or proportion of individuals within a population in various states (“compartments”) using a set of differential equations. Examples of compartmental models include the susceptible-infected-recovered (SIR) model [8], susceptible-infected-susceptible (SIS) model [9], and susceptible-infected-recovered-susceptible (SIRS) model [10], which are used to describe diseases spreading where recovery from infection provides permanent immunity, no immunity, and time-limited immunity, respectively. For diseases that have long latent periods in spreading, like seasonal flu, the susceptible-exposed-infected-recovered (SEIR) model [11] is used instead. For a detailed study on how to obtain $R_0$ for various compartmental models, see [12].

The compartmental models described above are based on mean-field approximations, i.e. the behavior of the population is considered representable by the mean behavior of all individuals in the same compartment. In reality, the transmission rates of some individuals (“superspreaders”) can be substantially higher than the average. Network-based models [13], [14] can be used to model the disease-spreading capability of individuals. Network models can also be used to describe the interactions between people and epidemics in multiple cities. It was shown [15] that restricting migration, while delaying the spread of disease between cities, did not necessarily constrict the epidemic peak in some cities.

Nevertheless, due to the simplicity of mean-field models compared to network-based models, in this article we use a mean-field-approximation-based compartmental model to track the evolution of COVID-19 in a community and show how such a simple model can still lead to new insights regarding disease control. We not only make the E compartment contagious, but also add a new H compartment to the aforementioned SEIR model to represent isolated or hospitalized cases and examine the effects of migration into and out of the affected community, as well as the effects of parameter changes corresponding to various government interventions. Our model is the first to consider simultaneously all concerned features to describe COVID-19, although each has been considered separately in some previous work [16].

B. RELATED STUDIES ON COVID-19

During the initial spread of COVID-19, some traditional and modified SEIR models were used to predict the genesis of the epidemic in Wuhan [17]. However, these models generally did not consider the asymptomatic transmission capabilities of COVID-19, corresponding to “exposed” individuals in the SEIR model. This led to a significant underestimation of the extent of the COVID-19 spread. This omission was generally corrected in later studies; for example, the study in [18] emphasizes the importance of early interventions to shield susceptibles from infection, rather than targeting infected cases alone, especially for diseases with asymptomatic transmission such as COVID-19.

1) TRANSMISSION DYNAMICS

It is crucial to understand the transmission dynamics of the COVID-19. With an accurate estimate of the basic reproduction number, governments can take effective actions against the pandemic. It was shown in [19] that the basic reproduction number of the COVID-19 is higher than SARS. In [20], the transmission dynamics and the geographical characteristics of the COVID-19 pandemic in Italy is studied and the basic reproduction number is estimated for different areas. In [21], the importance of air pollution-to-human transmission and the human-to-human transmission is demonstrated, suggesting that the former is stronger than the latter. Furthermore, the containment measures in Italy in [22], where it is found that the sequence of restrictions imposed to mobility and human-to-human interactions can significantly reduce the virus transmission.

2) MIGRATION EFFECTS

The consideration of migration effects was also a key feature in many COVID-19 studies. In [23], the study on the early dynamics of transmission in Wuhan showed that newly introduced cases in one area might eventually lead to new outbreaks. In [24], it warns that a global pandemic might happen unless substantial interventions are taken globally. In [25], a modified SEIR model is used, with migration into and out of the susceptible and exposed states, showing a good fit between the estimated and observed data for three Chinese provinces and for China as a whole. Additionally, an artificial intelligence technique was used for model prediction, using training data obtained from SARS, and demonstrated to be remarkably accurate. Migration data for China was used in [26], predicting that the numbers of infections in most cities in China would peak between middle February and early March 2020, which was indeed the case. In light of these results, we will also incorporate migration into our proposed compartmental model.

3) TRAVEL RESTRICTIONS

With regards to the implementation of travel restrictions, in [27] a complex network model is implemented and two strategies are compared: adaptive clustering, which mimics
self-isolation of small cliques within a larger community, and instant clustering, which mimics the imposition of hard border controls between cities and/or countries. It was demonstrated that the adaptive clustering strategy was more effective for preventing epidemic spread than instant clustering, due to the remaining weak connections between clusters caused by the impossibility of perfect border control. Additionally, in [28] the adverse effect of delaying the introduction of travel restrictions is demonstrated: by the time border controls were imposed on Wuhan on 23 January 2020, many Chinese cities had already received a large number of infected travelers. While a slight delay in epidemic progression was observed abroad, it was concluded that travel bans are only effective when combined with a significant reduction in local disease transmission. On the other hand, a study of human outflow from Wuhan in the weeks before the 23 Jan 2020 lockdown [29] suggests a strong correlation between the migration flow size from Wuhan to each Chinese province and the scale of the epidemic in that province (up to mid-February 2020).

C. CONTRIBUTIONS OF THIS ARTICLE

Enlightened by the SEIR model and recognizing the features of COVID-19, in this article we propose a compartmental model to study the spreading process of COVID-19 and its impact on the public healthcare systems. We modify the SEIR model as follows: First, we modify the E compartment of the SEIR model to allow for asymptomatic transmissions; Second, we introduce a hospitalized (denoted H) compartment, thus forming an SEIHR model. Unlike the “infected” (I) compartment, individuals in the H compartment are assumed not to transmit the disease to others, due to strict quarantine measures within the hospital setting. As a result, we find that in the SEIHR model, the key factor influencing the overall transmission rate is that of the “exposed” (i.e. asymptomatic but potentially infectious) individuals, rather than those in the I compartment which are assumed to be transferred to the H compartment quickly. In other words, the quick transfer of infected people from the I compartment to the H compartment can effectively isolate these people from the susceptible population and further lower the speed of disease transmission. Note that the number of people in the E and I compartments are assumed to be unknown, and only people in the H compartment are observable as having the disease.

In addition to the H compartment, we also model the effect of migration to and from each of the various compartments in the SEIHR model. We introduce a parameter to describe the relative scale of the epidemic in the external region, and discover that, when the local pandemic is more severe than the global one, border control policies may not be very helpful for lowering the transmission speed of the virus. However, if we do not impose the border control policy, the exported cases will spread the disease worldwide. Next, we fit the parameters of our model to COVID-19 caseload data from five global regions and compare the quality of our fit against the SIR and SEIR models, as well as quantifying the effect of various local intervention measures. The results demonstrate that by using a single parameter to represent the overall strength of all local interventions at a given stage of the epidemic, we can capture the local dynamical changes and further forecast the epidemic trend.

Compared with the traditional SIR and SEIR model, the additional compartment (H) and the dynamics between all compartments help us capturing the property of the COVID-19 pandemic more accurately. The contribution presented here distinguishes itself from previous work on compartmental models by being able to more efficiently represent characteristics of infection-spread dynamics and interventions. This is reflected in the fitting accuracy achieved for a given control effort. Furthermore, the proposed model is heuristically justifiable and interpretable from a human perspective, requiring only five compartments, making it of great practical value in applied epidemiology.

II. MODEL AND ANALYSIS

With regards to the insight previously gained regarding the COVID-19 pandemic, we propose an SEIHR model with migration to estimate the progression of COVID-19 outbreak. In addition to modeling human migration, our proposed SEIHR model differs from the traditional SEIR model in that we add a hospitalized (H) compartment and make the exposed (E) state contagious, to model transmissions from both asymptomatic and pre-symptomatic people. The proposed model thus consists of five states: susceptible (S), exposed (E), infected (I), hospitalized (H) and removed (R). The number of deaths in the proposed SEIHR model is estimated as a fixed proportion of the number of removed people. Natural births and deaths can be incorporated as “migration” to and from a non-alive state.

Table 1 shows all the parameters used in the SEIHR model. The system of differential equations that describes the SEIHR model can be written as

$$\frac{dS(t)}{dt} = \left[ -\varepsilon k \frac{E(t)}{N(t)} - \alpha k \frac{I(t)}{N(t)} + (\Delta_{in} - \Delta_{out}) \right] S(t), \quad (1a)$$

Table 1. Definition of parameters. Note that not all parameters are used in some of the models in this article.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon$</td>
<td>Transmission rate of exposed people</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Transmission rate of infected people</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Rate of becoming symptomatic after the latent period</td>
</tr>
<tr>
<td>$\delta_E$</td>
<td>Self-recovery rate of exposed people without hospitalization</td>
</tr>
<tr>
<td>$\delta_I$</td>
<td>Self-recovery rate of infected people without hospitalization</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Hospitalization rate of infected people</td>
</tr>
<tr>
<td>$\delta_H$</td>
<td>Recovery rate of hospitalized people</td>
</tr>
<tr>
<td>$\Delta_{in}$</td>
<td>Migration rate into the system, per capita</td>
</tr>
<tr>
<td>$\Delta_{out}$</td>
<td>Migration rate out of the system, per capita</td>
</tr>
<tr>
<td>$k$</td>
<td>Average number of people that each individual encounters daily</td>
</tr>
<tr>
<td>$p$</td>
<td>Local intervention index</td>
</tr>
<tr>
<td>$q$</td>
<td>External epidemic strength</td>
</tr>
</tbody>
</table>
Assuming that a state change does not occur first, people in one of these compartments (S and E) will leave the population at a rate of $\Delta_{\text{out}}$. Meanwhile, for people in the S or E compartment, new people will enter that compartment from outside the system at a rate of $\Delta_{\text{in}}$. In other words, the migration rate into and out of each compartment is proportional to the size of that compartment. On the other hand, the H compartment does not allow immigration or emigration. Finally, the I compartment allows both immigration and emigration; however, all migrants into or out of the I compartment are immediately detected and transferred to the H compartment, i.e. “emigrants” from the I compartment are not actually allowed to leave the community.

### A. BASIC REPRODUCTION NUMBER

It is important to determine the basic reproduction number, denoted $R_0$, of the disease propagation process. To do so, define $X = (S, E, I, H, R)$. Assume that there is a disease-free equilibrium (DFE), namely $E^0 = (S, 0, 0, 0, 0)$, in system (1a–1e). We begin by deriving the next-generation matrix [12] of our proposed system (1a–1e). First, we define the matrices $F$, denoting the rate of new individuals into the infectious E and I compartments, and $V$, denoting the rate of transfer of individuals from the E and I compartments to the non-infectious H and R compartments. Matrices $F$ and $V$ can be written respectively as

$$F = \begin{bmatrix} k\varepsilon + \Delta_{\text{in}} & k\alpha \\ 0 & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} \beta + \delta_E + \Delta_{\text{out}} & 0 \\ -\beta & \gamma + \delta_I + \Delta_{\text{out}} \end{bmatrix}.$$

The next-generation matrix of the system is defined as $FV^{-1}$:

$$FV^{-1} = \begin{bmatrix} 1 & 0 \\ \frac{\beta + \delta_E + \Delta_{\text{out}}}{\beta} & 0 \\ \frac{k\varepsilon + \Delta_{\text{in}}}{k\alpha} & 1 \\ \frac{\beta + \delta_E + \Delta_{\text{out}}}{\beta} & \frac{\gamma + \delta_I + \Delta_{\text{out}}}{\beta} \\ \frac{(\beta + \delta_E + \Delta_{\text{out}})(\gamma + \delta_I + \Delta_{\text{out}})}{0} & 0 \end{bmatrix}.$$

According to [12], the basic reproduction number $R_0$ is the largest eigenvalue or spectral radius of the next-generation matrix. Since the eigenvalues of a triangular matrix are the elements of its main diagonal,

$$R_0 = \rho(FV^{-1}) = \frac{k\varepsilon + \Delta_{\text{in}}}{\beta + \delta_E + \Delta_{\text{out}}} + \frac{k\alpha}{(\beta + \delta_E + \Delta_{\text{out}})(\gamma + \delta_I + \Delta_{\text{out}})}.$$

\[ \text{FIGURE 1. Inter-compartamental dynamics of the SEIHR model.} \]
The first term in $R_0$ is the number of people who are infected by other exposed people during one infection cycle, and the second term is the number of people who are infected by other infected people. When $R_0 > 1$, the disease outbreak will be sustained; when $R_0 < 1$, the disease will die out.

B. ACCOUNTING FOR EXTERNAL EPIDEMIC STRENGTH

According to Fig. 1, when modeling migration, the strength of the epidemic in the external population is assumed to be equal to the strength of the epidemic in the local population. However, in many situations, this is not the case; for example, during the initial stages of the COVID-19 pandemic, the number of infected individuals in Wuhan was much higher than in the rest of the world. Therefore, we introduce a scaling factor of $q \geq 0$ to denote the external epidemic strength, which represents the ratio of the percentage of individuals in the external regions that are exposed or infected compared to the local region. Then, system (1a–1e) can be written as

\[
\begin{align*}
\frac{dS(t)}{dt} &= \left[ -\varepsilon k \frac{E(t)}{N(t)} - \alpha k \frac{I(t)}{N(t)} \right] S(t) + \left[ \left( 1 - \frac{(q-1)(E(t)+I(t))}{S(t)} \right) \Delta_{in} - \Delta_{out} \right] S(t), \\
\frac{dE(t)}{dt} &= \left[ \varepsilon k \frac{E(t)}{N(t)} + \alpha k \frac{I(t)}{N(t)} \right] S(t) + [\beta - \delta_E + q \Delta_{in} - \Delta_{out}] E(t), \\
\frac{dI(t)}{dt} &= \beta E(t) + [-\gamma - \delta_I - \Delta_{out}] I(t), \\
\frac{dH(t)}{dt} &= \gamma + q \Delta_{in} + \Delta_{out} I(t) - \delta_H H(t), \\
\frac{dR(t)}{dt} &= \delta_E E(t) + \delta_I I(t) + \delta_H H(t). 
\end{align*}
\]

Similarly to Sec. 2.1, the basic reproduction number $R_0$ of system (3a–3e) is derived as

\[
R_0 = \frac{k E + q \Delta_{in}}{\beta + \delta_E + \Delta_{out}} + \frac{k \alpha \beta}{(\beta + \delta_E + \Delta_{out})(\gamma + \delta_I + \Delta_{out})}.
\]

If $q > 1$, it means the global pandemic is more severe than the local epidemic. When $q = 1$, the global pandemic has the same infection level as the local epidemic. While for $q < 1$, the local epidemic is more severe than in other regions. Note that for systems without immigration, i.e. $\Delta_{in} = 0$, the evolution of the system is independent of $q$ and one can assign the value of $q = 1$ for model fitting purposes.

C. SENSITIVITY OF $R_0$ TO THE MODEL PARAMETERS

In carrying out simulations, some basic assumptions are made. The latent period is set as 7 days ($\beta = 0.14$), the hospitalization rate $\gamma$ is set as 1, and the recovery rate is set as $\delta_E = \delta_I = \delta_H = 0.1$.

Figure 2 shows how $R_0$ changes according to changes in the various model parameters. Figure 2a shows that the
transmission rate of exposed individuals, \( \varepsilon \), has a major influence on \( R_0 \), compared with the transmission rate of infected individuals, \( \alpha \). The reason why this happens is because the hospitalization rate \( \gamma \) is very high, meaning each infected individual does not have much time to infect others. Figure 2b shows that social distancing (when \( k \) is small) can effectively reduce \( R_0 \) for highly contagious strains (when \( \varepsilon = \alpha \) is large). From Fig. 2c, without border control, when the transmission rate is low enough, the epidemic will not emerge even the scale of the external epidemic is larger than the local one. According to Fig. 2d, border control (when \( \Delta \) is small) may make the local epidemic be even worse when the scale of the external epidemic \( (q < 2) \) is not very large. Essentially, this is because a high liquidity of people can dilute the density of the exposed and infected individuals.

There are other ways to reduce \( R_0 \), such as finding a method to lower the transmission rate or increasing the recovery rate of the infected individuals. Due to the assumption that hospitalized individuals will not be contagious, the rate \( R_0 \) is not affected by the hospitalized individuals. However, the recovery rate of hospitalized individuals are influenced by the occupancy of medical facilities. If the recovery rate is too low, it will increase hospital occupancy and ultimately overflow the hospitals. Therefore, it is crucial to expand the capacity of the hospitals during the pandemic.

III. SIMULATION RESULTS AND DISCUSSIONS

A. NUMERICAL SIMULATION

In order to study the spreading process of COVID-19, we generated a regular network of size \( N = 10^6 \). By using the Runge-Kutta method, we simulated varies scenarios of the epidemic process. Initially, we set \( S(0) = 10^6 - 1, E(0) = 1, I(0) = 0, H(0) = 0, \) and \( R(0) = 0 \).

First, we investigated the impact of the transmission rate \( \varepsilon \) and parameter \( \alpha \). As shown by Fig. 3, for a large transmission rate, the peak of the epidemic comes very early, while for a smaller transmission rate the peak will be postponed. One can also see that, when the transmission rate \( \varepsilon \) is small, the size of the epidemic is much smaller compared with the results shown in Figs. 3b and 3c. Furthermore, given a fixed transmission rate \( \varepsilon \), the evolution of the epidemic is nearly insensitive to changes in \( \alpha \). This is consistent with the results shown in Fig. 2a. In other words, the epidemic process is more sensitive to the transmission rate of exposed people, \( \varepsilon \).

The influence of the patients on medical facilities is studied by changing the rate \( \gamma \) of infected people being hospitalized. Figure 4 shows the number of people being hospitalized to evolve over time under different hospitalization rates. The peak \( H_{\text{max}} \) of \( H(t) \) and the time when it arrives are related with \( \gamma \). When any of these peaks is larger than the local medical facility capacity, the hospitals will be overwhelmed. It can also be seen that the peaks of \( H(t) \) have a maximum value.

Next, we simulated the epidemic process under two different border control policies, where \( \Delta_{\text{in}} = \Delta_{\text{out}} = 0 \) represents total lock down and \( \Delta_{\text{in}} = \Delta_{\text{out}} = 0.1 \) represents normal operation. Figure 5 shows that the policy is continuously implemented from day one to the day epidemic ends. It can be seen that the spreading speed of the disease under both conditions are almost the same. This result can also be derived
Fig. 7. For each value of $\Delta$ and $q$, there exists a critical value $p_c$, such that when $p > p_c$, the scale of the epidemic increases rapidly, while for $p < p_c$, the outbreak is totally suppressed. This is simply because when $p = p_c$, the effective reproduction number equals 1.

Furthermore, when $q = 1$, which means the local pandemic is equivalent in strength to the global one, the critical value $p_c$ is similar for both the open-border and closed-border cases. On the other hand, when $q < 1$, i.e. the local pandemic is more severe than the global one, the critical value $p_c$ is larger than in the closed-border case, meaning that in terms of controlling the local epidemic, less severe interventions are required when the borders remain open. On the other hand, maintaining open borders when $q < 1$ comes at the cost of worsening the global pandemic due to exported cases.

B. REAL-DATA ANALYSIS

We used the proposed SEIHR model to fit the real COVID-19 data chosen from five representative regions, i.e., Italy, Germany, Florida, New York and Hong Kong, to validate the proposed model. Note that all the cases calculated here are from natural growth data, two basic transmission measures are implemented, an index $p$ is introduced to denote the strength of various interventions, including social distancing, environmental disinfection, wearing masks, and so on, to control the disease. An intervention strength of 0 means that the transmission rates of exposed and infected persons are both scaled by a factor of 1. Furthermore, we introduce a scaling factor of $\delta$ to denote the strength of various interventions, including social distancing, environmental disinfection, wearing masks, and so on, to control the disease. An intervention strength of 0 means that the transmission rates of exposed and infected persons are both scaled by a factor of 1. The black solid line, red solid line and blue solid line represent the cumulative number of cases for different transmission rates.

from Fig. 2c, since their $R_0$ values are very close to each other. However, the final numbers of these cases are smaller when the border is totally closed. This is because the infected individuals from outside are also counted when the border remains open.

Finally, we investigated what kinds of intervention measures would work for epidemic control. To do so, we first need to quantify the intervention measures. As in Fig. 5, we consider two possible states for border control, i.e. totally closed ($\Delta = \Delta_{in} = \Delta_{out} = 0$) and remaining open ($\Delta = \Delta_{in} = \Delta_{out} = 0.1$). Furthermore, we introduce a scaling factor of $p$ to denote the strength of various interventions, including social distancing, environmental disinfection, wearing masks, and so on, to control the disease. An intervention strength of 0 means that the transmission rates of exposed and infected persons are both scaled by a factor of 1, becoming $pE$ and $pI$, respectively. The updated dynamical process between compartments can be seen in Fig. 6. A $p$ value of one denotes no intervention, whereas a $p$ value of zero denotes a complete cessation of local transmission. Thus, we can further obtain the effective reproduction number as

$$R_e = \frac{p\alpha + q\Delta_{in}}{\beta + \delta_E + \Delta_{out}} + \frac{p\alpha\beta}{(\beta + \delta_E + \Delta_{out})(\gamma + \delta_I + \Delta_{out})}. \quad (5)$$

Equation (5) shows a way to fit the epidemic process to empirical data by tuning parameters $p$ and $q$.

One can see a threshold for the intervention strength from Fig. 7. For each value of $\Delta$ and $q$, there exists a critical value $p_c$, such that when $p > p_c$, the scale of the epidemic increases rapidly, while for $p < p_c$, the outbreak is totally suppressed. This is simply because when $p = p_c$, the effective reproduction number equals 1.
Inter-compartmental dynamics of the SEIHR model when parameters $p$ and $q$ are introduced.

![FIGURE 6]  

Relationship between the final number of cumulative case and $p$, where $k = 1$, $\beta = 0.14$, $\gamma = 1$, $\delta_E = \delta_I = \delta_H = 0.1$, $\epsilon = 0.25$, and $\alpha = 0.5$. The red solid line represent the final number of cumulative cases for different levels of intervention with or without closing the border.

![FIGURE 7]  

The SEIHR model fits with the second wave of COVID-19 outbreak in Hong Kong. The colored solid lines and dots represent simulation results and real data, respectively.

![FIGURE 8]  

The SEIHR model prediction for the third wave of COVID-19 outbreak in Hong Kong. The colored solid lines and dots represent simulation results and real data, respectively.

migration data is shown in Fig. 13), which fits the observed number of total cases very well.

However, for the third wave of COVID-19 outbreak in Hong Kong, the situation is slightly different. The number of imported cases was minimal due to strict border control. On the other hand, the number of local cases increased rapidly because of loosening interventions. In Fig. 9, we successfully predict the trend of the third wave of COVID-19 in Hong Kong using the SEIHR model without migrations. The parameter choices for both waves are shown in Table 5.

2) OTHER REGIONS
Figure 10 shows the COVID-19 data in four regions and the corresponding SEIHR simulation results. As shown in Fig. 10, the SEIHR model fits well with the real data. For comparison, we also show the fitting results of SIR and SEIR models. The results demonstrate that our modified SEIHR model is much more accurate than the SIR and SEIR models.

For Italy, we assume that the intervention period has only one stage which lasts to the end of the pandemic. In reality, different interventions were implemented in different stages. Therefore, the fitted curve for the SEIHR model deviates from the real data.

For Germany and Florida, we assume two stages of intervention, with overall intervention strengths of $p_1$ and $p_2$, respectively. Additionally, $p_1 < p_2$, i.e. in the second stage, certain interventions were relaxed. It can be seen that once the intervention measures were loosened, the infection rates in both regions increased immediately. This is because, $R_e$ will grow when intervention measures are reduced.

Note that for Italy, Germany, and Florida, the number and timing of each intervention period in our modified SEIHR...
models were estimated. For New York, to make the fit more reasonable, we consider the actual dates when certain intervention policies were implemented. We selected four intervention policies that might have large impacts on the pandemic, and divided the intervention period into four stages accordingly, with intervention strengths of $p_1$, $p_2$, $p_3$ and $p_4$, respectively. For the first three stages, each stage is more strict than the previous stage. In contrast, the fourth stage corresponds to a loosening of intervention policies. Therefore, $p_1 < p_2 < p_3 < p_4$. Figure 10d shows that the model fits the real data quite well.

3) DISCUSSION
As can be seen from the results, our modified SEIHR model achieves greater fitting accuracy compared with the SIR and SEIR models. This suggests that our choice of dynamical interaction between compartments is better able to capture hidden unidentified and even unidentifiable dynamics in this extremely non-linear and time-varying complex environment. Furthermore, by keeping all parameters except $p$ and $q$ constant throughout the course of an epidemic, our model requires fewer parameters, updated less frequently, than many other epidemiological models, e.g. network-based models. The simulation and data-based analysis results not only validate the proposed modified SEIHR model but also present a way to predict the future evolution of COVID-19 in general.

According to our data analysis, we now know that the intervention measures must be strictly and continuously enforced and implemented.

IV. LIMITATIONS OF THE STUDY
First, this study is built on an assumption that the human population is mixed homogeneously. However, in reality, the population distributions are mostly heterogeneous. The heterogeneity of the social contact network sometimes causes the emergence of super-spreaders, which play an important role in the epidemic process.

Second, the model we used is a deterministic approach for modeling the COVID-19 pandemic. As shown in the real data, there are some stochastic phenomena during the pandemic. Therefore, our deterministic approach cannot fully model the daily number of cases of an epidemic, as demonstrated in Fig. 10b. Nevertheless, numerical results show that our deterministic approach is still able to provide an accurate estimate of the total number of cumulative cases in a region.

In summary, while our model is much simpler than other models (deterministic with only five compartments), and does not model certain known phenomena such as heterogeneity or the stochastic nature of disease transmission, it can still obtain relatively accurate predictions for the COVID-19 pandemic.

V. CONCLUDING REMARKS
In this article, we presented a deterministic framework referred to as the SEIHR compartmental model. Using the next-generation matrix method, we derived an explicit expression of the basic reproduction number $R_0$. After
analyzing the relationship between $R_0$ and the model parameters, we obtained some parameter regions that allow to control the epidemic outbreak. We also found that, when the local pandemic is more severe than the global one, allowing people to move freely into or out of the system can actually reduce the speed of epidemic, at the price that it will export diseases to other areas. Then, we performed several simulations, showing that the epidemic process is more sensitive to the transmission rate of the exposed people than that of the infected people, due to a strict isolation policy for infected people who exhibit symptoms. Furthermore, we investigated how the isolation policy impacts the medical facilities and found a possible parameter region to lower the risk of hospital overflow. To that end, we used an intervention index to quantify the strengths of some measures implemented for local epidemic prevention. Our results show that, when the local epidemic is more severe than other regions, hard intervention measures for epidemic control could be more effective than blindly closing the borders. In addition, by using a set of real historical COVID-19 data, we validated the model and found that reopening a region too early may cause another wave of pandemic.

Finally, we verified that our model can estimate the evolution of local epidemics (with or without migration) by fitting the real data of the second and third waves of COVID-19 outbreak in Hong Kong, as well as COVID-19 outbreaks in four other global regions. These numerical results for multiple global regions demonstrate that our proposed modified SEIHR model is more accurate and robust than the SIR and SEIR models for estimating COVID-19 caseloads. In particular, our model can accurately model the number of both local and imported cases during the second wave of COVID-19 in Hong Kong, using a single parameter to model the relative strength of the COVID-19 pandemic outside of Hong Kong.

**APPENDIX A**

**ADDITIONAL FIGURES**

Figure 11 shows the relationship between $H_{\text{max}}$ and $\gamma$, where $k = 1$, $\varepsilon = 0.25$, $\alpha = 0.5$, $\beta = 0.14$, $\delta = \delta_M = 0.1$ and $\Delta_{\text{in}} = \Delta_{\text{out}} = 0$. The black squares represent $H_{\text{max}}$ for different values of $\gamma$.

Figure 12. Contour graph of the critical value $\gamma_c$ with respect to $\varepsilon$ and $\alpha$, where $k = 1$, $\beta = 0.14$, $\delta = \delta_M = 0.1$ and $\Delta_{\text{in}} = \Delta_{\text{out}} = 0$. The contour graph represents different values of $\gamma_c$.

Figure 13. Evolution of the SEIHR model when intervention measures are involved, where $q = 1$, $k = 1$, $\beta = 0.14$, $\gamma = 1$, $\delta = \delta_M = 0.1$, $\varepsilon = 0.25$, and $\alpha = 0.5$. The black solid line, red solid line, blue solid line and pink solid line represent the cumulative number of cases for different levels of intervention without closing the border. The green solid line represents the cumulative number of cases without intervention, but with border closed.

Figure 14. The migration data of Hong Kong during the second wave of the COVID-19 pandemic.

(a) Daily migration data

(b) Migration rate
strict isolation policies are implemented, the peak $H_{\text{max}}$ will decrease. At this time, the scale of the epidemic will also be reduced.

Figure 12 shows how the critical value $\gamma_c$ is changed with $\varepsilon$ and $\alpha$. As shown in the figure, the contour graph is divided into three regions. In region I, $\gamma_c = 1$, which means that when $\gamma \rightarrow 1$ the peak $H_{\text{max}}$ of $H(t)$ will increase. However, due to small values of $\varepsilon$ and $\alpha$, the value of the peak $H_{\text{max}}$ is small. In region II, $\gamma_c = 1$ is also reached. Under this circumstance, the peak $H_{\text{max}}$ of $H(t)$ will increase and its value is much higher due to the large values of $\varepsilon$ and $\alpha$. In region III, because $\gamma_c < 1$, when $\gamma \rightarrow 1$ the peak $H_{\text{max}}$ of $H(t)$ decreases, which further lowers the risk of hospital overflow.

Figure 13 shows the evolution of the SEIHHR model when intervention measures are involved. The intervention is taken when the number of cumulative cases exceeds 100.

Figure 14 shows the migration data of Hong Kong between February 28th and April 30th.

APPENDIX B
FITTED PARAMETER VALUES FOR THE FIVE GLOBAL REGIONS

Tables 2–5 contain the fitted parameter values for the five global regions analyzed in Section III-B, namely Italy, Germany, Florida, New York, and Hong Kong. For Hong Kong, three different predictions are given for the third wave of the COVID-19 epidemic, using data up to July 24, Aug 2, and Aug 23, respectively.

REFERENCES


**RUIWU NIU** received the B.Sc. and M.Sc. degrees in theoretical physics from Hubei University, Wuhan, China, in 2012 and 2016, respectively, and the Ph.D. degree from the School of Mathematics and Statistics, Wuhan University, Wuhan, in 2019. He is currently a Research Assistant with the Department of Electronic Engineering, City University of Hong Kong.

**ERIC W. W. M. WONG** (Senior Member, IEEE) received the B.Sc. and M.Phil. degrees in electronic engineering from The Chinese University of Hong Kong, Hong Kong, in 1988 and 1990, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Massachusetts at Amherst, Amherst, MA, USA, in 1994. He is currently an Associate Professor with the Department of Electrical Engineering, City University of Hong Kong, Hong Kong. His research interests include analysis and design of telecommunications and computer networks, energy-efficient data center design, green cellular networks, and optical networking.

**GUANRONG CHEN** (Life Fellow, IEEE) received the M.Sc. degree in computer science from Sun Yat-sen University, Guangzhou, China, in 1981, and the Ph.D. degree in applied mathematics from Texas A&M University at College Station, College Station, TX, USA, in 1987. He has been a Chair Professor and the Founding Director of the Centre for Chaos and Complex Networks, City University of Hong Kong, since 2000. Prior to that, he was a tenured Full Professor with the University of Houston, TX, USA. He is a member of the Academy of Europe and a Fellow of The World Academy of Sciences. He received the State Natural Science Award of China, in 2008, 2012, and 2016. He was awarded the 2011 Euler Gold Medal, Russia, and conferred Honorary Doctorates by the Saint Petersburg State University, Russia, in 2011, and the University of Le Havre, Normandy, France, in 2014. He has been a Highly Cited Researcher in Engineering according to Thomson Reuters, since 2009.

**YIN-CHI CHAN** (Member, IEEE) received the B.Math. degree from the University of Waterloo, Waterloo, ON, Canada, in 2010, and the M.Sc. and Ph.D. degrees from the City University of Hong Kong, Hong Kong, in 2011 and 2017, respectively. He is currently a Research Associate with the Department of Electronic Engineering, City University of Hong Kong. His research interests include performance evaluation, optimization, and design of communications and service systems.

**MICHAËL ANTONIE VAN WYK** (Senior Member, IEEE) received the B.Eng. and M.Eng. degrees in electrical engineering and the Ph.D. degree in applied mathematics from Rand Afrikaans University, Johannesburg, South Africa, in 1986, 1989, and 1996, respectively. From 1987 to 1992, he completed his compulsory national service and worked in the aerospace industry. In 1992, he joined Rand Afrikaans University as a Lecturer. He is currently a Full Professor and the Carl and Emily Fuchs Chair in Systems and Control with the School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, and the Head of the Systems and Control Group. He has recently been appointed as an Adjunct Professor at the Department of Electrical Engineering, Nelson Mandela University, Port Elizabeth, South Africa. He has been a consultant in the aerospace and electronic systems industry in the areas of signal processing and control, nationally and internationally. His current research interests include analysis and modeling of nonlinear dynamical systems and networks and algorithm development in signal and image processing.