Research article

# Deterministic and stochastic models for the epidemic dynamics of COVID-19 in Wuhan, China

**Damilola Olabode**[†]**, Jordan Culp**[†]**, Allison Fisher, Angela Tower, Dylan Hull-Nye and Xueying Wang**[*]

Department of Mathematics and Statistics, Washington State University, Pullman, WA 99164, USA

† First authors with equal contribution.

* **Correspondence:** Email: xueying@math.wsu.edu; Tel: +15093356996.

**Abstract:** In this paper, deterministic and stochastic models are proposed to study the transmission dynamics of the Coronavirus Disease 2019 (COVID-19) in Wuhan, China. The deterministic model is formulated by a system of ordinary differential equations (ODEs) that is built upon the classical SEIR framework. The stochastic model is formulated by a continuous-time Markov chain (CTMC) that is derived based on the ODE model with constant parameters. The nonlinear CTMC model is approximated by a multitype branching process to obtain an analytical estimate for the probability of a disease outbreak. The local and global dynamics of the disease are analyzed by using the deterministic model with constant parameters, and the result indicates that the basic reproduction number $\mathcal{R}_0$ serves as a sharp disease threshold: the disease dies out if $\mathcal{R}_0 \leq 1$ and persists if $\mathcal{R}_0 > 1$. In contrast to the deterministic dynamics, the stochastic dynamics indicate that the disease may not persist when $\mathcal{R}_0 > 1$. Parameter estimation and validation are performed to fit our ODE model to the public reported data. Our result indicates that both the exposed and infected classes play an important role in shaping the epidemic dynamics of COVID-19 in Wuhan, China. In addition, numerical simulations indicate that a second wave of the ongoing pandemic is likely to occur if the prevention and control strategies are not implemented properly.

**Keywords:** COVID-19; SEIR model; ODE model; CTMC model; disease extinction, disease outbreak; second wave

## 1. Introduction

On December 31, 2019, a highly contagious pneumonia case with unknown origin was reported in Wuhan City, Hubei Province of China. Wuhan has been the focus of worldwide attention due to

the outbreak of Coronavirus Disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The number of infections and deaths has been growing rapidly since then. COVID-19 has become a pandemic [1] and disease is creating unprecedented public health challenges worldwide. As of August 30, 2020, the disease has caused over 12.5 million confirmed cases and 843 thousand deaths in at least 188 countries and territories.

COVID-19 is a new disease. The origin and etiology of the infection is still uncertain. There were likely substantial proportions of undetected cases in early periods [2–4]. Moreover, the disease incubation period ranges from 2 to 14 days [5]. During this period of time, infected individuals may not show any symptoms; thus, they may not have awareness of the infection, but they are capable of transmitting the disease to other people. Therefore, we are facing unprecedented challenges in the prevention and control of the disease. To that end, a large number of mathematical models have being proposed to improve our understanding in mitigating the pandemic (see, e.g., [1, 3, 6–20] and the references therein). Some of these works are summarized below. Chen et al. [13] cataloged the symptoms and physical characteristics of the first 425 laboratory-confirmed cases in Wuhan. Their study shows that the basic reproduction number $\mathcal{R}_0$ is 2.2 and the incubation period of these early cases to be 7.5 ± 3.4 days. Wu et al. [4] presented a susceptible-exposed-infectious-recovered (SEIR) model in which the reported value of $\mathcal{R}_0$ is 2.68 and they estimated that the COVID-19 epidemic was doubling every 6.4 days. Their study helped to forecast the potential domestic and international spread of COVID-19. Ndairou et al. [18] used a compartmental epidemic model to study the spread of the COVID-19 with an inclusion of super spreaders. In their case study of Wuhan, the reported value of the basic reproduction number is 0.945, indicating that the containment within Wuhan was well-controlled by the Chinese authorities. A modeling study in [4] inferred the basic reproduction number, outbreak size in Wuhan, exported cases from Wuhan, forecasting the spread in China. In [21], a modified SEIR model and an AI trained model on SARS 2003 data were proposed to predict the epidemic curve and promise for future prediction of epidemics. Lin et al. [22] used a variation of the SEIR model to study the outbreak of COVID-19 in Wuhan by incorporating individual behavioral change and governmental actions towards the spread of the disease. Hao et al. [23] employed a more detailed epidemic model to help understand the vigorous non-pharmaceutical intervention that has helped in keeping infection spread at the lowest in Wuhan. However, most of these mathematical models of COVID-19 dynamics are deterministic. To capture the discrete nature of human population and heterogeneity among individual hosts, a stochastic modeling approach may be more appropriate, particularly during the initiation of the infection. That motivates our work.

The goal of this article is to develop deterministic and stochastic models to study the evolution of the disease in both short and long terms. Our focus is mathematical modeling and analysis of disease epidemic, which aims to address the important questions such as *how long will it take to disease extinction and how likely the second wave of the pandemic would occur if the reopening of Wuhan is not well contained?*

The remainder of this paper is organized as follows. Section 2 presents the deterministic ordinary differential equation (ODE) model and performs a detailed analysis on the local and global dynamics of the ODE model with constant parameters. Section 3 is dedicated to the formulation of the stochastic model and a theoretical estimate of the probability of disease extinction. In Section 4, parameter estimation is carried out to fit the publicly reported data and numerical simulations of deterministic and stochastic models are performed. Finally we conclude the paper with some discussion in Section 5.

## 2. Deterministic model

To study the transmission dynamics of the COVID-19 epidemic, we employ the classical SEIR epidemic modeling framework where the total human population is composed of four compartments: the susceptible, the exposed, the infected and the recovered. Let $S, E, I, R$ denote the number of human individuals in each of the classes; thus, our deterministic model takes the form:

$$
\begin{aligned}
\frac{dS}{dt} &= b - \beta(t)I\frac{S}{N} - dS, \\
\frac{dE}{dt} &= \beta(t)I\frac{S}{N} - (d + \alpha)E, \\
\frac{dI}{dt} &= \alpha E - (d + \tau + \gamma_I)I, \\
\frac{dR}{dt} &= \tau I - dR.
\end{aligned}
\tag{2.1}
$$

Here $N = N(t)$ is the total population size at time $t$. The parameter $b$ and $d$ are the natural birth and death rates, respectively. Let $N_0 = b/d$ denote the initial population size. The function $\beta(t)$ is the transmission rate. This function is time dependent and assumed to be piecewise constant, which is to capture the dramatic change in the prevention and control policy in Wuhan, such as city lockdown, hospitalization and quarantine. Parameter $\alpha^{-1}$ denotes the latent period, $\tau^{-1}$ is the (mean) infectious period and $\gamma_I$ is the disease-induced mortality rate. For simplicity of analysis, in the rest of this section, we assume that $\beta(t) = \beta$ is constant.

### 2.1. Basic reproduction number $\mathcal{R}_0$

The basic reproduction number $\mathcal{R}_0$ in epidemic models is an important threshold parameter to understand the extinction and persistence of a disease. It is defined as the expected number of secondary infectious cases produced by a primary infected individual in an otherwise susceptible population. Using the next generation matrix method [24], we compute $\mathcal{R}_0$ and it is given by

$$
\mathcal{R}_0 = \frac{\alpha\beta}{(d + \alpha)(d + \tau + \gamma_I)}.
$$

### 2.2. Equilibrium analysis

It is easy to verify that the deterministic model (2.1) admits at most two biologically feasible equilibrium solutions which depends on the value of $\mathcal{R}_0$.

The disease-free equilibrium (DFE) $\mathcal{E}_0 = (N_0, 0, 0, 0)$ always exists and $\mathcal{E}_0$ is the only feasible equilibrium when $\mathcal{R}_0 \leq 1$. An endemic equilibrium (EE)

$$
\mathcal{E}_1 = (S^*, E^*, I^*, R^*) = \left( N_0 - \frac{1}{\mathcal{R}_0}\frac{\beta}{d}I^*, \ \frac{d + \tau + \gamma_I}{\alpha}I^*, \ I^*, \ \frac{\tau}{d}I^* \right)
\tag{2.2}
$$

for which $I^* = \frac{b}{\beta - \gamma_I}(\mathcal{R}_0 - 1)$ exists when $\mathcal{R}_0 > 1$. Note that $\mathcal{R}_0 > 1$ implies $\beta > \gamma_I$.

### 2.2.1. Local stability

The local stability of the DFE is an immediate consequence of [24, Theorem 2]. More specifically, the DFE is locally asymptotically stable for $\mathcal{R}_0 < 1$ and unstable for $\mathcal{R}_0 > 1$. The following theorem establishes the local stability of the EE of model (2.1).

**Theorem 2.1.** *The EE, $\mathcal{E}_1$, of the system* (2.1) *is locally asymptotically stable for $\mathcal{R}_0 > 1$.*

*Proof.* Let $S^*, E^*, I^*, R^*$ be as defined above as in (2.2). Note that the Jacobian matrix of the linearized system of (2.1) at the EE $\mathcal{E}_1$ is

$$J = \begin{pmatrix} -\frac{\beta I^*}{N^*} + \frac{\beta I^* S^*}{(N^*)^2} - d & \frac{\beta I^* S^*}{(N^*)^2} & -\frac{\beta S^*}{N^*} + \frac{\beta I^* S^*}{(N^*)^2} & \frac{\beta I^* S^*}{(N^*)^2} \\ \frac{\beta I^*}{N^*} - \frac{\beta I^* S^*}{(N^*)^2} & -\frac{\beta I^* S^*}{(N^*)^2} - (d + \alpha) & \frac{\beta S^*}{N^*} - \frac{\beta I^* S^*}{(N^*)^2} & -\frac{\beta I^* S^*}{(N^*)^2} \\ 0 & \alpha & -(d + \tau + \gamma_I) & 0 \\ 0 & 0 & \tau & -d \end{pmatrix}.$$

Hence the corresponding characteristic equation is

$$(\lambda + d)(\lambda^3 + a_1\lambda^2 + a_2\lambda + a_3) = 0,$$

where

$$a_1 = \alpha + d + d + \gamma_I + \tau + \frac{b}{S^*},$$

$$a_2 = \frac{b}{S^*}(\alpha + d + d + \gamma_I + \tau),$$

$$a_3 = \frac{\alpha\beta b}{\mathcal{R}_0 N^*}(\mathcal{R}_0 - 1).$$

By the Routh-Hurwitz criterion, the EE is locally stable when $a_1 > 0$, $a_3 > 0$ and $a_1 a_2 > a_3$. Clearly $a_1 > 0$, since $\alpha, d, \tau, b, \gamma_I$, and $S^* > 0$. Similarly, $a_3 > 0$, for $\alpha, \beta, b, N^* > 0$ and $\mathcal{R}_0 > 1$. Now it remains to show that $a_1 a_2 - a_3 > 0$. Let $u = \alpha + d$ and $v = d + \tau + \gamma_I$. Thus,

$$\begin{aligned} a_1 a_2 - a_3 &= \left(u + v + \frac{b}{S^*}\right)\left(\frac{b}{S^*}(u + v)\right) - uv\frac{b}{S^*}\left(1 - \frac{1}{\mathcal{R}_0}\right) \\ &= \frac{b}{S^*}(u + v)^2 + \left(\frac{b}{S^*}\right)^2(u + v) - uv\frac{b}{S^*}\left(1 - \frac{1}{\mathcal{R}_0}\right) \\ &= \frac{b}{S^*}(u^2 + v^2) + \left(\frac{b}{S^*}\right)^2(u + v) + \frac{b}{S^*}uv\left(1 + \frac{1}{\mathcal{R}_0}\right). \end{aligned}$$

This shows that $a_1 a_2 - a_3 > 0$.
Thus, the local stability result of the EE follows when $\mathcal{R}_0 > 1$. $\qquad\square$

### 2.2.2. Global stability

It is clear that the biologically feasible region

$$\Omega = \{(S, E, I, R) \in \mathbb{R}_+^4 : S + E + I + R \leq N_0\}$$

is the positive invariant for model (2.1). The global stability of the DFE and the EE is established in the following two theorems.

**Theorem 2.2.** *If $\mathcal{R}_0 \leq 1$, then the DFE $\mathcal{E}_0$ is globally asymptotically stable in $\Omega$.*

*Proof.* One can verify that

$$\begin{bmatrix} \frac{dE}{dt} \\ \frac{dI}{dt} \end{bmatrix} \leq (F - V) \begin{bmatrix} E \\ I \end{bmatrix}$$

where

$$F = \begin{bmatrix} 0 & \beta \\ 0 & 0 \end{bmatrix}, V = \begin{bmatrix} d + \alpha & 0 \\ -\alpha & d + \gamma_I + \tau \end{bmatrix}.$$

Define $\mathbf{x} = \begin{bmatrix} E & I \end{bmatrix}^T$, and $\mathbf{u} = (0, \beta)$. Consider a Lyapunov function $L = \mathbf{u}V^{-1}\mathbf{x}$. Taking the derivative of $L$ along the solution of model (2.1) in $\Omega$ leads to

$$\dot{L} := \frac{dL}{dt} = \mathbf{u}V^{-1}\frac{d\mathbf{x}}{dt} \leq \mathbf{u}V^{-1}(F - V)\mathbf{x} \leq \mathbf{u}(V^{-1}F - I)\mathbf{x}.$$

Note that $\mathbf{u}$ is the left eigenvector of the matrix $V^{-1}F$ corresponding to the spectral radius of $V^{-1}F$. Since $\mathcal{R}_0 = \rho(FV^{-1}) = \rho(V^{-1}F)$, $\mathbf{u}V^{-1}F = \mathcal{R}_0\mathbf{u}$. Thus,

$$\dot{L} \leq \mathbf{u}(\mathcal{R}_0 - 1)\mathbf{x}$$

If $\mathcal{R}_0 < 1$, $\dot{L} \leq 0$ and $\dot{L} = 0$ implies that $\mathbf{u}\mathbf{x} = 0$. This shows that $I = 0$. Using the second, fourth and the first equations of model (2.1), we have $E = 0$, $R = 0$ and $S = N_0$. This shows that the largest invariant set on which $\dot{L} = 0$ is the singleton $\mathcal{E}_0$.

If $\mathcal{R}_0 = 1$, $\dot{L} = 0$ and the fact $\mathbf{u}V^{-1} > 0$ yield $S = N$. By the definition of $N = S + E + I + R$ and $E, I, R \geq 0$, $E = I = R = 0$. In view of the first equation of model (2.1), $S = N_0 = b/d$. Hence, the largest invariant set on which $\dot{L} = 0$ is the singleton $\mathcal{E}_0$. Therefore, by LaSalle's invariance principle [25], the DFE is globally asymptotically stable in $\Omega$ when $\mathcal{R}_0 \leq 1$. $\qquad\square$

**Theorem 2.3.** *If $\mathcal{R}_0 > 1$ and $\gamma_I = 0$, then the unique endemic equilibrium $E_1$ is globally asymptotically stable in the interior of $\Omega$.*

*Proof.* By $\gamma_I = 0$, $dN/dt = b - dN$. Since $N(0) = N_0 = b/d$, $N(t) \equiv N_0$ for all $t \geq 0$. Recall that $\mathcal{E}_1 = (S^*, E^*, I^*, R^*)$. Consider a Lyapunov function $V$ as follows:

$$V = \left(S - S^* - S^* \ln \frac{S}{S^*}\right) + \left(E - E^* - E^* \ln \frac{E}{E^*}\right) + c\left(I - I^* - I^* \ln \frac{I}{I^*}\right) \tag{2.3}$$

where $c = \frac{\beta S^* I^*}{\alpha E^* N}$. It is clear that $V \geq 0$ and $V = 0$ if and only if $U = U^*$ for $U = S, E, I$. Taking the derivative of (2.3) along the solutions of model (2.1), we have the following,

$$\dot{V} := \frac{dV}{dt} = \dot{S}\left(1 - \frac{S^*}{S}\right) + \dot{E}\left(1 - \frac{E^*}{E}\right) + c\dot{I}\left(1 - \frac{I^*}{I}\right)$$

Using the equilibrium equations of model (2.1) yields

$$\dot{V} = \left(b - \beta\frac{S}{N}I - dS\right)\left(1 - \frac{S^*}{S}\right) + \left(\beta\frac{S}{N}I - (d + \alpha)E\right)\left(1 - \frac{E^*}{E}\right)$$

$$+ c(\alpha E - (d + \tau + \gamma_I)I)\left(1 - \frac{I^*}{I}\right)$$

$$= dS^*\left(2 - \frac{S}{S^*} - \frac{S^*}{S}\right) + \beta\frac{S^*}{N^*}I^*\left(3 - \frac{S^*}{S} - \frac{I^*}{I}\frac{E}{E^*} - \frac{S}{S^*}\frac{I}{I^*}\frac{E}{E^*}\right)$$

By the arithmetic-geometric mean inequality,

$$\frac{S}{S^*} + \frac{S^*}{S} \geq 2, \qquad \frac{S}{S^*} + \frac{I^*E}{IE^*} + \frac{S}{S^*}\frac{I}{I^*}\frac{E}{E^*} \geq 3$$

and the quantities hold if and only if $S = S^*$, $E = mE^*$ and $I = mI^*$ for some $m > 0$. Substituting $I = mI^*$ into the fourth equation of model (2.1) yields $R = mR^*$. Using the first equations leads to $m = 1$. Hence the largest invariant set where $\dot{V} = 0$ is the singleton $\mathcal{E}_1 = (S^*, E^*, I^*, R^*)$. It follows from the LaSalle's Invariance Principle [25], the EE $\mathcal{E}_1$ is globally asymptotically stable in the interior of $\Omega$ when $\mathcal{R}_0 > 1$.

$\square$

Theorems 2.2 and 2.3 show that $\mathcal{R}_0$ serves as a sharp threshold for disease dynamics. More specifically, the disease will die out if $\mathcal{R}_0 < 1$ and persist if $\mathcal{R}_0 > 1$.

## 3. Stochastic process

Epidemic ODE models typically assume that the sizes of the compartments are large enough that the mixing of members is homogeneous, or at least that there is homogeneous mixing in each subgroup if the population is stratified by activity levels. However, at the beginning of a disease outbreak, there is a very small number of infectious individuals and the transmission of infection is indeed a stochastic event depending on heterogeneity among individuals in the population (e.g., variations in individual health conditions and disease transmissibility) and patterns of contacts between them. This suggests that the homogeneous mixing at the beginning of an epidemic may not be a good assumption and hence the ODE models may not be appropriate when the initial infection is small. To that end, we use a continuous time Markov chain (CTMC) model to study the variability of the disease dynamic, and then we apply the theory of the multitype branching process (e.g., see [26–28] and references therein) to approximate the dynamics of the CTMC model near the DFE.

### 3.1. Continuous time Markov chain (CTMC) model

We use the ODE model (2.1) with constant parameters as a framework to formulate a CTMC model, which is composed of nine independent events. For simplicity, we use the same notation as in the ODE model (2.1). Let $\mathbf{Z}(t) = (S(t), E(t), I(t), R(t))$ be a discrete random vector and $\Delta\mathbf{Z}(t) = \mathbf{Z}(t + \Delta t) - \mathbf{Z}(t)$ denote the change during $[t, t + \Delta t]$, for $t \in [0, \infty)$. The infinitesimal transition probabilities are defined based on the ODE rates. The changes and the corresponding transition probabilities are summarized in Table 1.

### 3.2. Branching process approximation (BP)

To investigate the dynamics of the CTMC model near the DFE (i.e., $S \approx N_0$), we employ the multitype branching process theory by approximating the "birth and death" of the infection process near the origin. Assume that $S = N_0$ and $R = 0$ and consider only the events in Table 2. The transition rates in this table follow directly from the rates in the linearized ODE model (2.1). This leads to a multitype branching process for the exposed and infected human individuals, where a transition occurs in one of the infectious random variables, $E$ or $I$.

**Table 1.** Transition probabilities of the CTMC model.

| Event | Description | $\Delta\mathbf{Z}(t)$ | Transition probability |
|---|---|---|---|
| 1 | Birth of $S$ | (1, 0, 0, 0) | $b\Delta t + o(\Delta t)$ |
| 2 | Death of $S$ | (-1, 0, 0, 0) | $dS(t)\Delta t + o(\Delta t)$ |
| 3 | Infection of $S$ | (-1, 1, 0, 0) | $\left(\beta\frac{S(t)}{N(t)}I(t)\right)\Delta t + o(\Delta t)$ |
| 4 | Natural death of $E$ | (0, -1, 0, 0) | $dE(t)\Delta t + o(\Delta t)$ |
| 5 | Incubation loss of $E$ | (0, -1, 1, 0) | $\alpha E(t)\Delta t + o(\Delta t)$ |
| 6 | Natural death of $I$ | (0, 0, -1, 0) | $dI(t)\Delta t + o(\Delta t)$ |
| 7 | Disease-induced death | (0, 0, -1, 0) | $\gamma_I I(t)\Delta t + o(\Delta t)$ |
| 8 | Recovery of $I$ | (0, 0, -1, 1) | $\tau I(t)\Delta t + o(\Delta t)$ |
| 9 | Natural death of $R$ | (0, 0, 0, -1) | $dR(t)\Delta t + o(\Delta t)$ |

**Table 2.** Transition probabilities $r_i\Delta t + o(\Delta t)$ for the branching process approximation, where $\Delta\mathbf{X}(t) = \mathbf{X}(t + \Delta t) - \mathbf{X}(t)$.

| Event | Description | $\Delta\mathbf{X}(t)$ | Rate $r_i$ |
|---|---|---|---|
| 1 | Infection | (1, 0) | $\beta I(t)$ |
| 2 | Natural death of $E$ | (-1, 0) | $dE(t)$ |
| 3 | Incubation loss of $E$ | (-1, 1) | $\alpha E(t)$ |
| 4 | Loss of $I$ | (-1, 0) | $(d + \tau + \gamma_I)I(t)$ |

In general, let $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ denote a vector of discrete-state random variables where $X_i$ is the random variable corresponding to the infectious group $i$ for $i = 1, 2, \cdots, n$. Let $X_i(0) = \delta_{ij}$ where $\delta_{ij}$ is the Kronecker delta with $\delta_{ij} = 1$ if $i = j$ and zero otherwise, for $1 \le i, j \le n$. The offspring pgf for the infectious individuals of type $i$ is a function $f_i : [0, 1]^n \to [0, 1]$ and

$$f_i(x_1, x_2, \cdots, x_n) = \sum_{k_1=0}^{\infty} \sum_{k_2=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} p_i(k_1, k_2, \cdots, k_n) x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$$

where $p_i(k_1, k_2, \cdots, k_n)$ denotes the probability that the individual of type $i$ gives "birth" to $k_j$ individuals of type $j$, for $j = 1, 2, \cdots, n$. In our case, $\mathbf{X} = (X_1, X_2) = (E, I)$ and $n = 2$.

By Table 2, the offspring pgf for $E$ is given $(E(0), I(0)) = (1, 0)$ is

$$f_1(x_1, x_2) = \frac{d + \alpha x_2}{d + \alpha}$$

and the offspring pgf for $I$ given $(E(0), I(0)) = (0, 1)$ is

$$f_2(x_1, x_2) = \frac{\beta x_1 x_2 + \tau + \gamma_I + d}{\beta + \tau + \gamma_I + d}.$$

It follows from branching process theory that the key of estimating the probability of disease extinction is to determine the fixed point of the offspring pgf on $[0, 1]$. Solving $f_i(q_1, q_2) = q_i$ for $i = 1, 2$ on $[0, 1]$ leads to

$$q_1 = \frac{d + \alpha q_2}{d + \alpha},$$
$$(q_2 - 1)(c_1 q_2 + c_2) = 0,$$

where $c_1 = \alpha\beta$ and $c_2 = -(d + \alpha)(\tau + \gamma_I + d)$.

Thus fixed points in terms of $q_2$ on $[0, 1]$ are 1 and $q_2^* = -c_2/c_1 = 1/\mathcal{R}_0$. It is clear that $0 < q_2^* < 1$ if and only if $\mathcal{R}_0 > 1$, which shows that the minimal fixed point of $q_2$ is $q_2^* = \min\{1/\mathcal{R}_0, 1\}$. Thus the fixed point $(q_1^*, q_2^*)$ of $f_i(q_1, q_2) = q_i$ $(i = 1, 2)$ in $[0, 1]^2 \backslash \{(1, 1)\}$ is given by

$$
\begin{aligned}
q_1^* &= \frac{d}{d + \alpha} + \frac{\alpha}{d + \alpha} \frac{1}{\mathcal{R}_0}, \\
q_2^* &= \frac{1}{\mathcal{R}_0},
\end{aligned}
\tag{3.1}
$$

when $\mathcal{R}_0 > 1$.

By the theory of the multitype branching process approximation, an estimate of the probability of disease extinction is given by

$$
\mathbb{P}_{ext} = \begin{cases} (q_1^*)^{e_0}(q_2^*)^{i_0}, & \text{if } \mathcal{R}_0 > 1, \\ 1, & \text{if } \mathcal{R}_0 \leq 1, \end{cases}
\tag{3.2}
$$

and consequently, the probability of an outbreak is

$$
\mathbb{P}_{out} = 1 - \mathbb{P}_{ext} = \begin{cases} 1 - \left((q_1^*)^{e_0}(q_2^*)^{i_0}\right), & \text{if } \mathcal{R}_0 > 1, \\ 0, & \text{if } \mathcal{R}_0 \leq 1, \end{cases}
\tag{3.3}
$$

where $e_0, i_0$ corresponds to the initial population size of the exposed class and the infected class respectively. This particularly shows that the outbreak probability is $1 - q_1^*$ (resp. $1 - q_2^*$) when the infection is initiated by one exposed (resp. infected) individual.

The above results highlight the difference between the deterministic and stochastic dynamics; i.e., unlike the deterministic dynamics showing that the disease persists and reaches an endemic equilibrium when $\mathcal{R}_0 > 1$, the stochastic model indicates there is a positive probability of disease extinction when $\mathcal{R}_0 > 1$.

## 4. Numerical simulations

In this section, we fit our ODE model to the public reported data, conduct a sensitivity analysis for the model parameters and predict the second wave of the pandemic in terms of the time, duration and strength. On the other hand, numerical simulations of the CTMC model are performed. In particular, we use the stochastic model to investigate the disease extinction and the first and second waves of the disease epidemics in terms of the time and the likelihood.

### 4.1. Deterministic model

#### 4.1.1. Parameter estimation

Using the method of iterated filtering [29], we fit our ODE model to COVID-19 data published daily by WHO and other sources [30–34] from December 31, 2019 to June 28, 2020. It has been estimated that the population of Wuhan was approximately 9,000,000 people [35], so that we set $N(0) = 9,000,000$ in our model. The model parameters are taken from [36–38], which is summarized in

**Table 3.** Definition and value of the parameters. Here the unit of time is a day.

| Parameter | Description | Value |
|-----------|-------------|-------|
| $\alpha^{-1}$ | latent period (days) | 2 |
| $\tau^{-1}$ | mean infectious period (days) | 5 |
| $\gamma_I$ | disease induced mortality rate | 1/17.5 |
| $d$ | crude death rate per 1000 people | 10.94 |

Table 3. The other initial compartment values are set as $S(0) = 8,999,768$, $E(0) = 39$, $I(0) = 193$, and $R(0) = 0$.

We assume that $\beta(t)$ is a piecewise constant function of time. We estimate $\beta(t)$ by fitting to our data sets in five time intervals: days 0–17, days 17–28, days 28–37, days 37–42, and days greater than 42. These days were chosen as follows. Day 17 corresponds to January 23, 2020 which is the day that the citywide lockdown began in Wuhan, China [39]; Day 28 corresponds to February 2, 2020 which is the day that the centralized quarantine and treatment began [39] (and also this is the end of Period 1 in [2]); Day 37 corresponds to the first day in which clinical diagnosis of COVID-19 were included in the dataset on February 12, 2020 [40]; Day 42 corresponds to February 17, 2020 which is the day that the community universal symptom survey began and is the peak in daily active infected cases from our calculations [39]. It is well-known that the dataset of confirmed cases for COVID-19 is greatly underestimated due to various reasons such as high transmissibility, long incubation periods, low levels of testing, particularly, in the beginning and high levels of asymptomatic infections (e.g., see [2,41–43] and the references therein). To account for under-reported cases, we assume an ascertainment rate of 14% before January 23, 2020 and then linearly increase the ascertainment rate from 14% to 65% until February 12, 2020 [2]. After February 12, 2020, the ascertainment rate is assumed to be 100% [44]. The cumulative infections in Wuhan estimated from our model are 3049 people by January 18, 2020 and cumulative cases are estimated at 36,091 people by January 27, 2020, which is in line with the results from other studies [22,43].
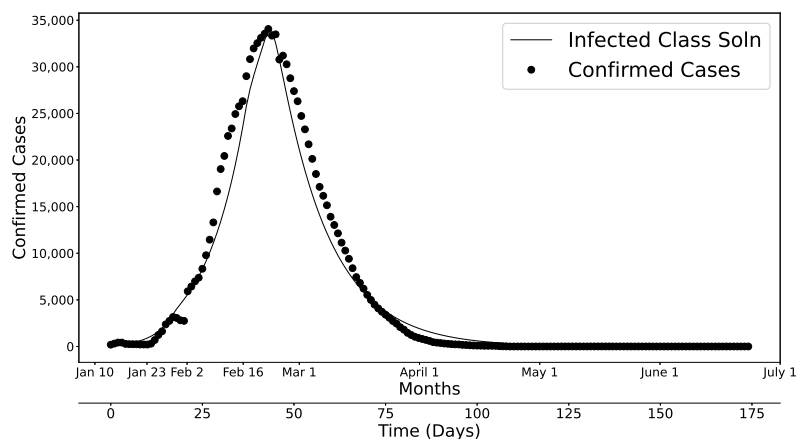


**Figure 1.** Fit to the confirmed cases for Wuhan from December 31, 2019 to June 28, 2020. Solid circles are the reported cases and the curve is the best fit to the data.

Figure 1 presents our fitted result to the data. It shows that (1) the onset of the epidemic occurs in January of 2020; (2) the epidemic lasts for about two months; (3) the situation of the epidemic is under control by March of 2020 and the disease tends to be extinct after July of 2020.
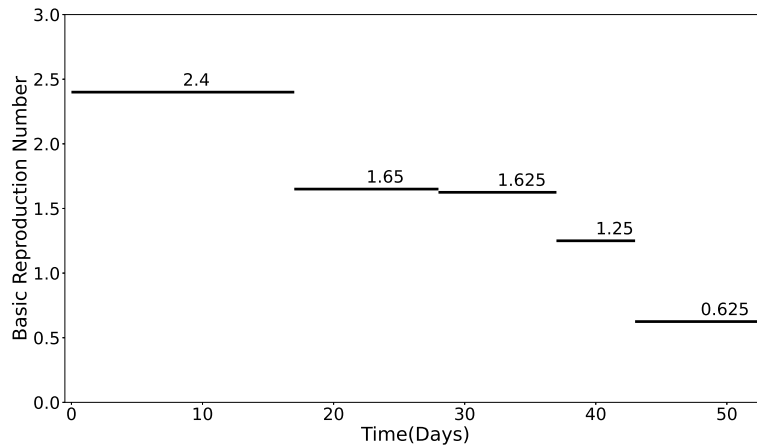


**Figure 2.** $\mathcal{R}_0(t)$ over the first 60 days.

Figure 2 plots the time-dependent effective reproduction number $\mathcal{R}_0(t)$ over the first 60 days, where

$$\mathcal{R}_0(t) = \frac{\alpha\beta(t)}{(d + \alpha)(d + \tau + \gamma_I)}.$$

This indicates that the effective reproduction number is 2.4 during $[0, 17)$ days, and it is gradually reduced to 1.25 by February 12, 2020 (Day 42) and and dropped to 0.625 afterward. The details of the estimated $\beta$ and $\mathcal{R}_0$ values are provided in Table 4.

**Table 4.** Estimated values of $\mathcal{R}_0(t)$ and $\beta(t)$.

| Parameter | 0–17 | 17–28 | 28–37 | 37–42 | > 42 |
|---|---|---|---|---|---|
| $\beta(t)$ | 0.617 | 0.424 | 0.418 | 0.321 | 0.161 |
| $\mathcal{R}_0(t)$ | 2.4 | 1.65 | 1.625 | 1.25 | 0.625 |

### 4.1.2. Prediction

As shown in Figure 3, the ODE model predicts the disease extinction after June of 2020, which is a direct consequence of the effective intervention, prevention and control strategies currently implemented in Wuhan. There arises a question: Will the second wave take place if these strategies can't be implemented properly, for instance due to economic reopening?

To address this question, we elevate the basic reproduction number $\mathcal{R}_0$ by increasing $\beta(t)$ after June 28, 2020 (175 days) at which $\mathcal{R}_0 = 0.625$. The simulated result of our ODE model is illustrated in Figure 3. Our fitted solution (based on the data until June 28, 2020) is displayed in black and the curves in blue (resp. orange, green, red) shows the evolution of the infected cases over time when the value of $\mathcal{R}_0$ is increased to 1.5 (resp. 2.0, 2.5, 3.0) after June 28, 2020. The prediction of our ODE

model shows that (a) the second wave may be much stronger than the first one; (b) the higher of the $\mathcal{R}_0$ value, the sooner the second wave would take place with higher peak value. This result indicates the second wave of the pandemic would happen if the infection risk goes up, for instance, due to changes in control strategy or human behavior.
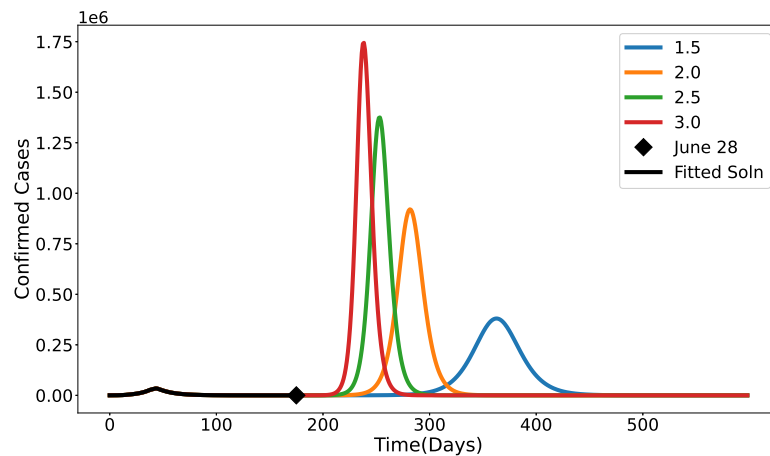


**Figure 3.** Impact of $\mathcal{R}_0$ on the second wave of the pandemic.

Additionally, in this section, we conduct numerical simulations using stochastic models.

First, we study the probability of a disease outbreak. The outbreak is assumed to be attained when the cumulative sum of $E$ and $I$ reaches $10,000$. The outbreak probability is estimated from the proportion of the sample paths (out of $10,000$) of the CTMC model. Then the obtained estimate is compared to the theoretical approximation computed by the multitype branching process theory. Figure 4 displays the probability of a disease outbreak as a function of $\mathcal{R}_0$ when the infection is initiated by an individual from the infected class ($I(0) = 1$), where $\mathcal{R}_0$ is changed by varying $\beta$. In the case where the infection is initiated by an exposed individual ($E(0) = 1$), the result is nearly identical to the case begun with an infected individual. This happens because of $d \ll \alpha$. Note that when $d \ll \alpha$, $\frac{d}{\alpha+d} \approx 0$ and $\frac{a}{a+d} \approx 1$ and hence it follows from (3.1) that $q_1^* = \frac{d}{d+\alpha} + \frac{\alpha}{d+\alpha}\frac{1}{\mathcal{R}_0} \approx \frac{1}{\mathcal{R}_0} = q_2^*$, which implies $1 - q_1^* \approx 1 - q_2^*$. This result indicates that the outbreak probability is (approximately) $1 - 1/\mathcal{R}_0$ when $\mathcal{R}_0 > 1$, which is regardless of the initial condition of the infection (beginning with $E(0) = 1$ or $I(0) = 1$), since the human lifespan ($1/d$) is much longer than the disease latent period ($1/\alpha$). In both cases, (1) we find a strong agreement between the theoretical estimate (using branching process theory) and the numerical simulation of the CTMC model (using Gillespie simulation algorithm); (2) the chance of an outbreak increases as $\mathcal{R}_0$ increases; (3) the disease extinction occurs with probability one (i.e., the probability of an outbreak is zero) when $\mathcal{R}_0 < 1$; (4) unlike the deterministic model that predicts the persistence of the disease, the stochastic model indicates there is a positive chance for disease extinction.

Secondly, we study the time to the disease extinction and an outbreak by using the CTMC models for 10,000 simulation runs in each scenario. Figure 5 displays the conditional probability distribution of the extinction time given the extinction takes place during $[0, 400]$ days, where the left (resp. right) panel displays the result when the infection starts with one exposed $E$ (resp. infected $I$) individual. It shows that there is a delay in the time to extinction when the infection is initiated by an exposed individual $E(0) = 1$ as compared to that by an infected individual $I(0) = 1$. Besides, in the case of $I(0) = 1$,
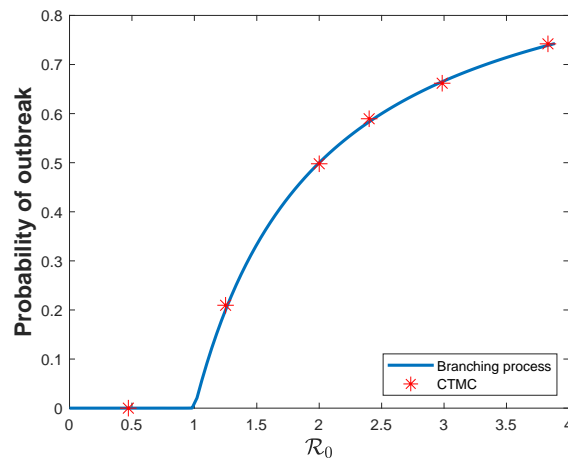
**Figure 4.** Probability of an outbreak as a function of $\mathcal{R}_0$, where $\mathcal{R}_0$ is changed by varying $\beta$. The solid curve is the analytical estimate obtained by the multitype branching process approximation (i.e., $1 - q_2^*$) and the stars are the result obtained from Monte Carlo simulation of the CTMC model for 10,000 simulation runs. The infection is initiated by one infected individual (i.e., $I(0) = 1$).
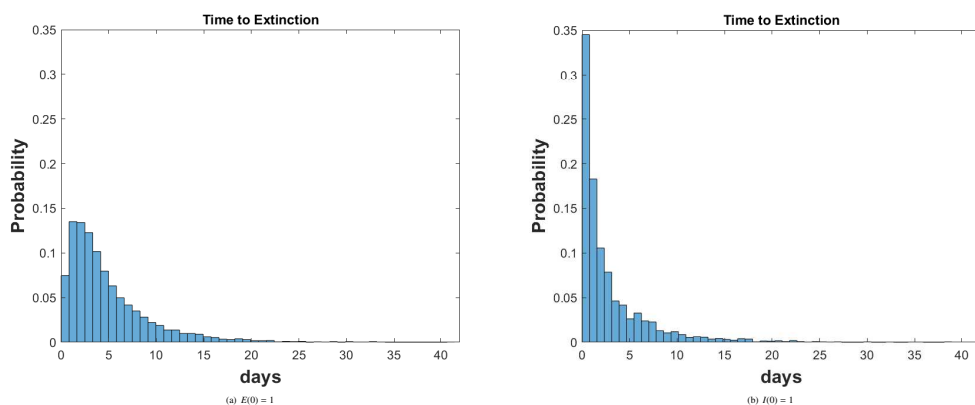


**Figure 5.** The probability distribution of the time to disease extinction given the occurrence of the extinction during $[0, 400]$ days when the infection is initiated by one individual in the (a) exposed class and (b) infected class.

the probability distribution of extinction time undergoes an immediate peak in the beginning and then declines rapidly as time increases. This indicates that disease extinction is expected to occur sooner with higher probability if the infection starts with one infected person. A biological interpretation of this result is that exposed individuals have to pass the incubation period to be capable of transmitting the disease to others and hence it delays the time to extinction. Figure 6 illustrates the conditional probability distribution of an outbreak given the occurrence of an outbreak during $[0, 400]$ days. If we fix a time during this investigation period, we see the that the outbreak probability (the corresponding area in Figure 6) is higher in the case where infection is initiated by a person from the infected class, as compared to that initiated with a person from the exposed class.
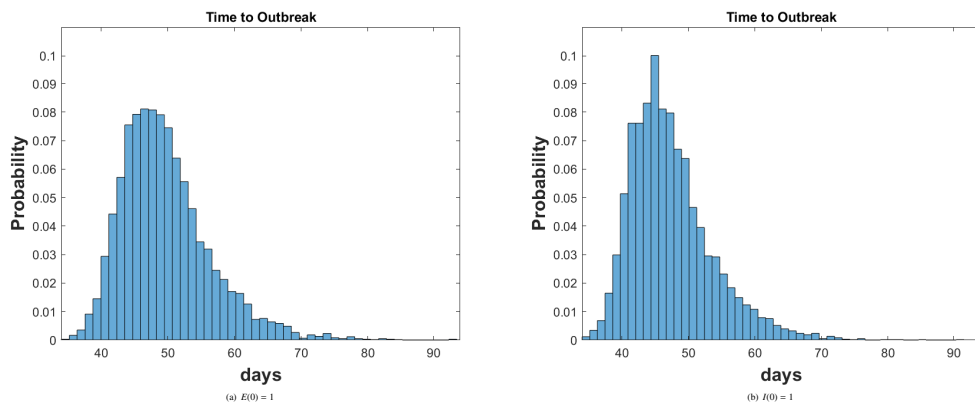
**Figure 6.** The probability distribution of the time to an outbreak given the occurrence of an outbreak during [0, 400] days when the infection is initiated by one individual in the (a) exposed class and (b) infected class.

**Table 5.** Mean and standard derivation of time to disease extinction and time to an outbreak for different $\mathcal{R}_0$ values given that the occurrence of an extinction and an outbreak during [0, 400] days.

| Initial condition | $\beta$ | $\mathcal{R}_0$ | $P_{out}$ | $E[T_{ext}|ext]$ | $\sigma[T_{ext}|ext]$ | $E[T_{out}|out]$ | $\sigma[T_{out}|out]$ |
|---|---|---|---|---|---|---|---|
| | 0.62 | 2.4 | 0.59 | 4.90 | 4.42 | 49.70 | 6.73 |
| | 0.51 | 2.0 | 0.50 | 5.72 | 5.60 | 64.39 | 8.68 |
| $E = 1, I = 0$ | 0.45 | 1.75 | 0.43 | 6.58 | 6.79 | 81.04 | 11.08 |
| | 0.39 | 1.5 | 0.34 | 8.03 | 9.11 | 113.57 | 16.11 |
| | 0.32 | 1.25 | 0.21 | 10.67 | 14.65 | 203.75 | 30.61 |
| | 0.62 | 2.4 | 0.59 | 2.96 | 4.04 | 47.38 | 6.18 |
| | 0.51 | 2.0 | 0.51 | 3.74 | 5.29 | 62.31 | 8.49 |
| $E = 0, I = 1$ | 0.45 | 1.75 | 0.43 | 4.62 | 6.57 | 79.28 | 11.04 |
| | 0.39 | 1.5 | 0.34 | 6.04 | 9.03 | 111.36 | 15.81 |
| | 0.32 | 1.25 | 0.20 | 9.05 | 15.08 | 201.85 | 30.23 |

Additionally, we investigate the time to the second wave of the pandemic. The current $\mathcal{R}_0$ is below one (from our investigation on the ODE model). Hence, to account for the change in restrictions and intervention and human behavior after reopening, we consider four different scenarios by assuming that $\mathcal{R}_0$ is elevated to 2.0, 1.75, 1.5 and 1.25. Figure 7 shows the conditional probability distribution of time to the second outbreak under these four scenarios when the infection starts with one $I$. Our result indicates that the second wave is likely to happen sooner with the higher value $\mathcal{R}_0$ after reopening. The mean and standard derivation for the time to extinction and the time to the second outbreak that are associated with the two initial conditions (i.e., either $E(0) = 1, I(0) = 0$ or $E(0) = 0, I(0) = 1$) are summarized in Table 5 for the selected $\mathcal{R}_0$ values. For instance, in the case where $\mathcal{R}_0$ is elevated to 1.5 after reopening, there would be about a 34% chance of a second outbreak and the standard deviation of the time to extinction is about 9 days. If we compare the infection initiated by an exposed individual to that initiated by an infected individual, as shown in Table 5, the average extinction time is reduced from 8.03 to 6.04 days, the average time to outbreak decreases from 113.57 to 111.36 days and the corresponding standard deviation is dropped slightly from 16.11 to 15.81 days. In these two scenarios, we see the difference in terms of time to extinction or outbreak is small. As $\mathcal{R}_0$ increases, both disease
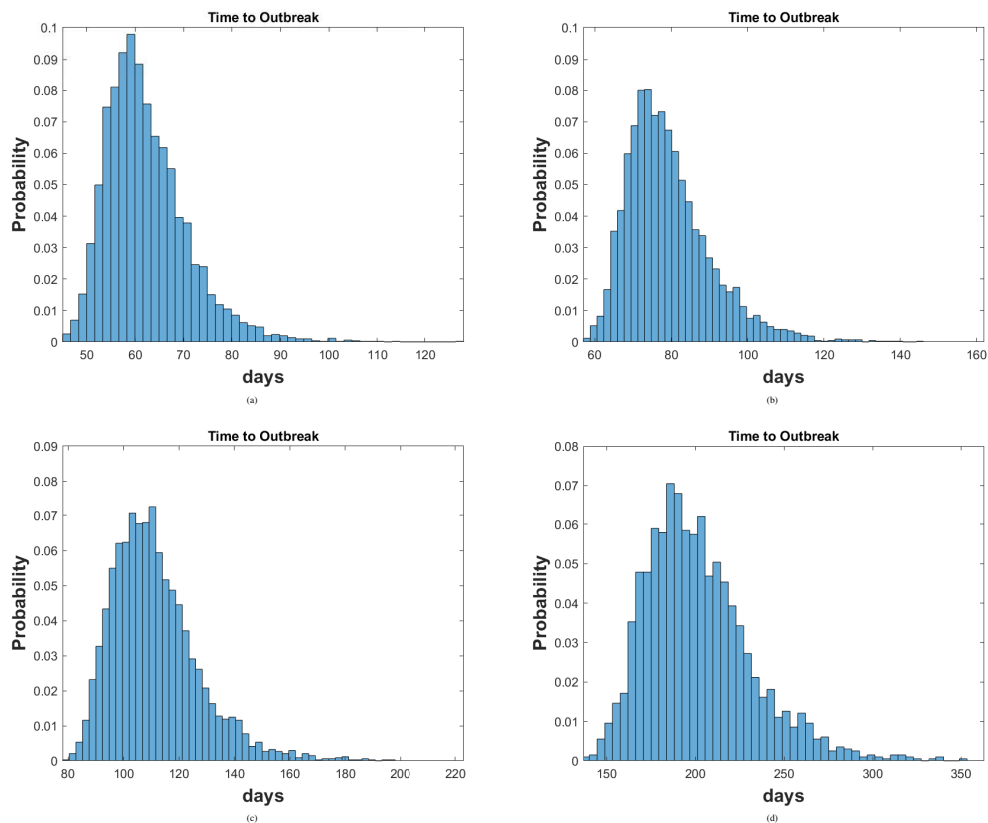
**Figure 7.** The probability distribution of the time to an outbreak given the occurrence of an outbreak during $[0, 400]$ days when the infection is initiated by one individual in the infected class for (a) $\mathcal{R}_0 = 2$, (b) $\mathcal{R}_0 = 1.75$, (c) $\mathcal{R}_0 = 1.5$, and (d) $\mathcal{R}_0 = 1.25$.

extinction and outbreak are likely to occur sooner. More specifically, if $\mathcal{R}_0$ is increased from 1.25 to 2.4, the average time to the second outbreak (resp. extinction) is decreased from 201.85 to 47.38 days (resp. from 9.05 to 2.96 days) in the case of the infection starting with an infected individual. Nevertheless, our results from deterministic and stochastic models indicate that there is a possibility of the second wave if the reopening is not handled properly.

## 5. Discussion

In this paper, we propose a deterministic and stochastic modeling framework to study the ongoing epidemic dynamics of the COVID-19 in the city of Wuhan, China. The deterministic model is formulated by a system of ODEs model that is built upon the classical SEIR framework. The stochastic model is formulated by a CTMC that is derived based on the ODE model with constant parameters during the early stage of the infection. We obtain a theoretical estimate for the probability of a disease outbreak by using multitype branching process approximation. Then we conduct a detailed mathematical analysis to the ODE model with constant model parameters, and our result indicates that the basic reproduction number $\mathcal{R}_0$ is served as a sharp disease threshold: the disease dies out if $\mathcal{R}_0 \leq 1$ and persists if $\mathcal{R}_0 > 1$. In contrast, the stochastic dynamics indicate that the disease may not persist when $\mathcal{R}_0 > 1$. A parameter estimation and validation is performed to fit our ODE model to the public reported

data that is corrected to address underestimation issue. Additionally, we use numerical simulations of the CTMC model to study the disease extinction and an outbreak. Our results show that if the non-pharmaceutical interventions such as quarantine and social distancing were not implemented properly, the second wave is likely to take place. There are several limitation of this work. First, our model is based on the classical SEIR epidemic framework. It would be more realistic if more compartments representing variations in the population (e.g., asymptomatic, symptomatic, hospitalized individuals and different age groups) are included. Second, it would be interesting to study the ongoing pandemic under various prevention and control strategies (e.g., the stay-at-home order, the reopening of schools and social distancing measures) with data available. Third, the epidemic patterns of COVID-19 vary significantly by countries and regions. The impact of spatial heterogeneity on the transmission and spread of COVID-19 will provide an interesting topic in future research.

## Acknowledgements

We thank the editor and reviewers for their helpful suggestions.

## Conflict of interests

The authors declare that they have no conflict of interests.

## Data availability

All data can be supplied by the authors upon request.

## References

1. M. V. Krishna, J. Prakash, Mathematical modelling on phase based transmissibility of corona virus, *Infect. Dis. Model.*, **5** (2020), 375–385.

2. R. Li, S. Pei, B. Chen, Y. Song, T. Zhang, W. Yang, et al., Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (sars-cov-2), *Science*, **368** (2020), 489–493.

3. A. J. Kucharski, T. W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, et al., Early dynamics of transmission and control of COVID-19: a mathematical modelling study, *Lancet Infect. Dis.*, **20** (2020), 553–558.

4. J. T. Wu, K. Leung, G. M. Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study, *Lancet*, **395** (2020), 689–697.

5. C. Rothe, M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, C. Wallrauch, et al., Transmission of 2019-ncov infection from an asymptomatic contact in germany, *N. Engl. J. Med.*, **382** (2020), 970–971.

6. C. Yang, J. Wang, A mathematical model for the novel coronavirus epidemic in Wuhan, China, *Math. Biosci. Eng.*, **17** (2020), 2708–2724.

7. L. Peng, W. Yang, D. Zhang, C. Zhuge, L. Hong, Epidemic analysis of COVID-19 in China by dynamical modeling, *arXiv preprint: 2002.06563*, 2020.

8. J. S. Weitz, S. J. Beckett, A. R. Coenen, D. Demory, M. Dominguez-Mirazo, J. Dushoff, et al., Modeling shield immunity to reduce COVID-19 epidemic spread, *Nat. Med.*, **26** (2020), 849–854.

9. P. Sookaromdee, V. Wiwanitkit, Imported cases of 2019-novel coronavirus (2019-ncov) infections in Thailand: Mathematical modelling of the outbreak, *Asian Pac. J. Trop. Med.*, **13** (2020), 139–140.

10. Z. Cakir, H. Savas, A mathematical modelling approach in the spread of the novel 2019 coronavirus sars-cov-2 (COVID-19) pandemic, *Electron. J. Gen. Med.*, **17** (2020), em205.

11. V. Volpert, M. Banerjee, S. Petrovskii, On a quarantine model of coronavirus infection and data analysis, *Math. Model. Nat. Phenom.*, **15**, 24.

12. P. Khrapov, A. Loginova, Mathematical modelling of the dynamics of the coronavirus COVID-19 epidemic development in China, *Int. J. Open Inf. Tech.*, **8** (2020), 13–16.

13. T.-M. Chen, J. Rui, Q.-P. Wang, Z.-Y. Zhao, J.-A. Cui, L. Yin, A mathematical model for simulating the phase-based transmissibility of a novel coronavirus, *Infect. Dis. Poverty*, **9** (2020), 1–8.

14. S. H. Khoshnaw, R. H. Salih, S. Sulaimany, Mathematical modelling for coronavirus disease (COVID-19) in predicting future behaviours and sensitivity analysis, *Math. Model. Nat. Phenom.*, **15** (2020), 33.

15. L. Zhong, L. Mu, J. Li, J. Wang, Z. Yin, D. Liu, Early prediction of the 2019 novel coronavirus outbreak in the mainland China based on simple mathematical model, *IEEE Access*, **8** (2020), 51761–51769.

16. S. He, S. Tang, L. Rong, A discrete stochastic model of the COVID-19 outbreak: Forecast and control, *Math. Biosci. Eng.*, **17** (2020), 2792–2804.

17. K. Chatterjee, K. Chatterjee, A. Kumar, S. Shankar, Healthcare impact of COVID-19 epidemic in India: A stochastic mathematical model, *Med. J. Armed Forces India*, **76** (2020), 147–155.

18. F. Ndairou, I. Area, J. J. Nieto, D. F. Torres, Mathematical modeling of COVID-19 transmission dynamics with a case study of Wuhan, *Chaos Soliton. Fract.*, **135** (2020), 109846.

19. S. He, Y. Peng, K. Sun, Seir modeling of the COVID-19 and its dynamics, *Nonlinear Dyn.*, **101** (2020), 1667–1680.

20. H. Zhao, Z. Feng, Staggered release policies for COVID-19 control: Costs and benefits of relaxing restrictions by age and risk, *Math. Biosci.*, **326** (2020), 108405.

21. Z. Yang, Z. Zeng, K. Wang, S.-S. Wong, W. Liang, M. Zanin, et al., Modified seir and ai prediction of the epidemics trend of COVID-19 in china under public health interventions, *J. Thorac. Dis.*, **12** (2020), 165–174.

22. Q. Lin, S. Zhao, D. Gao, Y. Lou, S. Yang, S. S. Musa, et al., A conceptual model for the outbreak of coronavirus disease 2019 (COVID-19) in Wuhan, China with individual reaction and governmental action, *Int. J. Infect. Dis.*, **93** (2020), 211–216.

23. X. Hao, S. Cheng, D. Wu, T. Wu, X. Lin, C. Wang, Reconstruction of the full transmission dynamics of COVID-19 in Wuhan, *Nature*, **584** (2020), 420–424.

24. P. Van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math. Biosci.*, **180** (2002), 29–48.

25. J. LaSalle, *The stability of dynamical systems*, Society for Industrial and Applied Mathematics, 1976.

26. L. J. Allen, G. E. Lahodny Jr, Extinction thresholds in deterministic and stochastic epidemic models, *J. Biol. Dyn.*, **6** (2012), 590–611.

27. K. B. Athreya, P. Jagers, *Classical and modern branching processes*, vol. 84, Springer Science & Business Media, 2012.

28. K. S. Dorman, J. S. Sinsheimer, K. Lange, In the garden of branching processes, *SIAM Rev.*, **46** (2004), 202–229.

29. E. L. Ionides, A. Bhadra, Y. Atchadé, A. King, Iterated filtering, *Ann. Stat.*, **39** (2011), 1776–1802.

30. W. H. Organization, WHO novel coronavirus (2019-nCoV) situation reports, 2020. Available from: `https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports`.

31. J. H. U. C. for Systems Science and Engineering, Covid-19 data repository, 2020. Available from: `https://github.com/CSSEGISandData/COVID-19`, Accessed: July 2020.

32. J. T. Wu, K. Leung, M. Bushman, N. Kishore, R. Niehus, P. M. de Salazar, et al., Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China, *Nat. Med.*, **26** (2020), 506–510.

33. N. H. Commission, National health commission of China daily reports on novel coronavirus, available from: `http://www.nhc.gov.cn/yjb/pzhgli/new_list.shtml`.

34. G. of Wuhan, The government of Wuhan homepage, available from: `http://english.wh.gov.cn`.

35. G. Huang, C. Che, R. Frost, Wuhan virus news: Coronavirus city lockdown of millions, 2020. Available from: `https://www.bloomberg.com/news/articles/2020-01-27/worried-angry-and-isolated-life-under--s-lockdown`.

36. R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, et al., Estimates of the severity of coronavirus disease 2019: a model-based analysis, *Lancet Infect. Dis.*, **20** (2020), 669–677.

37. J. H. University, Johns Hopkins University: New Study on Coronavirus Estimates 5.1 Days for Incubation Period - ProQuest, available from: `https://search.proquest.com/docview/2375478043`.

38. W. H. Organization, Report of the WHO-China joint mission on Coronavirus disease 2019, 16-24 February 2020, available from: `http://www.nhc.gov.cn/yjb/s2907/new_list.shtml`.

39. A. Pan, L. Liu, C. Wang, H. Guo, X. Hao, Q. Wang, et al., Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China, 2020. *Jama*, **323** (2020), 1915–1923.

40. T. K. Tsang, P. Wu, Y. L. Y. Lin, E. Lau, G. M. Leung, B. J. Cowling, Impact of changing case definitions for COVID-19 on the epidemic curve and transmission parameters in mainland China, *medRxiv*, (2020).

41. M. He, L. Li, L. P. Dehner, L. Dunn, Cremation based estimates suggest significant under-and delayed reporting of COVID-19 epidemic data in Wuhan and China, *medRxiv*, (2020).

42. C. Cadell, Data suggest virus infections underreported, exaggerating fatality rate, 2020. available from: `https://fr.reuters.com/article/us-china-health-deaths-idUKKBN1ZZ1AH`.

43. N. Imai, I. Dorigatti, A. Cori, C. Donnelly, S. Riley, N. Ferguson, Report 2: Estimating the potential total number of novel Coronavirus cases in Wuhan City, China, *Technical Report 2*, Imperial College London COVID-19 Response Team, London, UK, 2020.

44. J. Ge, D. He, Z. Lin, H. Zhu, Z. Zhuang, Four-tier response system and spatial propagation of COVID-19 in China by a network model, *Math. Biosci.*, **330** (2020), 108484.