Infectious Disease Modelling 6 (2021) 643–663

Contents lists available at ScienceDirect

Infectious Disease Modelling

journal homepage: www.keaipublishing.com/idm

Nonpharmaceutical interventions contribute to the control of COVID-19 in China based on a pairwise model

Xiao-Feng Luo [a, 1], Shanshan Feng [a, 1], Junyuan Yang [b, c], Xiao-Long Peng [b, c], Xiaochun Cao [b, c], Juping Zhang [b, c], Meiping Yao [d], Huaiping Zhu [e], Michael Y. Li [f], Hao Wang [f], Zhen Jin [b, c, *]

[a] Department of Mathematics, North University of China, Taiyuan, Shanxi, 030051, China
[b] Complex System Research Center, Shanxi University, Taiyuan, 030006, Shanxi, China
[c] Shanxi Key Laboratory of Mathematical Techniques and Big Data Analysis on Disease Control and Prevention, Shanxi University, Taiyuan, 030006, Shanxi, China
[d] School of Mathematical Sciences, Shanxi University, Taiyuan, 030006, Shanxi, PR China
[e] Department of Mathematics and Statistics, York University, Toronto, ON, M3J 1P3, Canada
[f] Department of Mathematics and Statistics Sciences, University of Alberta, Edmonton, Alberta, T6G 2G1, Canada

ARTICLE INFO

Article history:
Received 29 January 2021
Received in revised form 1 April 2021
Accepted 2 April 2021
Available online 10 April 2021

Keywords:
COVID-19
Pairwise epidemic model
Household quarantine
Clustering coefficient
High-risk contacts

ABSTRACT

Nonpharmaceutical interventions (NPIs), particularly contact tracing isolation and household quarantine, play a vital role in effectively bringing the Coronavirus Disease 2019 (COVID-19) under control in China. The pairwise model, has an inherent advantage in characterizing those two NPIs than the classical well-mixed models. Therefore, in this paper, we devised a pairwise epidemic model with NPIs to analyze COVID-19 outbreak in China by using confirmed cases during February 3rd–22nd, 2020. By explicitly incorporating contact tracing isolation and family clusters caused by household quarantine, our model provided a good fit to the trajectory of COVID-19 infections. We calculated the reproduction number $R = 1.345$ (95% CI: 1.230 – 1.460) for Hubei province and $R = 1.217$ (95% CI: 1.207 – 1.227) for China (except Hubei). We also estimated the peak time of infections, the epidemic duration and the final size, which are basically consistent with real observation. We indicated by simulation that the traced high-risk contacts from incubated to susceptible decrease under NPIs, regardless of infected cases. The sensitivity analysis showed that reducing the exposure of the susceptible and increasing the clustering coefficient bolster COVID-19 control. With the enforcement of household quarantine, the reproduction number R and the epidemic prevalence declined effectively. Furthermore, we obtained the resumption time of work and production in China (except Hubei) on 10th March and in Hubei at the end of April 2020, respectively, which is broadly in line with the actual time. Our results may provide some potential lessons from China on the control of COVID-19 for other parts of the world.

* Corresponding author. Complex System Research Center, Shanxi University, Taiyuan, 030006, Shanxi, China.
E-mail address: jinzhn@263.net (Z. Jin).
Peer review under responsibility of KeAi Communications Co., Ltd.
[1] The first two authors are contributed equally.

https://doi.org/10.1016/j.idm.2021.04.001

1. Introduction

The coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was firstly reported in Wuhan, a crowded city in central China's Hubei province (Li et al., 2020). Its symptoms mainly include cough, fever and difficult breath, similar to SARS (severe acute respiratory syndrome conronarirus) and MERS (middle east respiratory syndrome coronarirus) (Wu et al., 2020). These symptoms from mild to severe respiratory infections lasts about 2–14 days. The disease could be transmitted by direct transmission, such as touching or shaking hands, and indirect transmission such as coughing, sneezing and contacting with virus-contaminated substances (Wu et al., 2020). As of 7th December 2020, it has spread to more than 212 countries, causing more than 66 million confirmed cases and 1.5 million deaths worldwide (World Health Organization (WHO)) Especially, the second wave of COVID-19 is recently breaking out in Europe and America, which causes more new infections and deaths (World Health Organization (WHO)). Although COVID-19 was serious at the outset in China, it has been effectively brought under control. Nonpharmaceutical interventions (NPIs), particularly contact tracing isolation and household quarantine, made a great contribution to controlling the disease. In order to provide some potential lessons from the experience of containing COVID-19 for other parts of world, in this paper, we analyze the transmission course of COVID-19 in China.

In general, NPIs, such as isolation and quarantine, are the most effective intervention measures for a newly infectious disease since without available vaccines and cure medicine. Intervention of contact tracing isolation and household quarantine play a role of great importance in curbing COVID-19 in China. On January 23rd, 2020, with surge of the confirmed cases, Chinese government took an extreme control strategy of lockdowning Wuhan city, prohibiting all the transportation in and out of Wuhan, and advised all the Chinese citizens from leaving their homes unless they need to go outdoors for daily necessities (household quarantine) (Health Commission of Hubei Province). For the suspected and infected cases, they were traced, and once anyone was confirmed, he or she and individuals who have contacted with them would be isolated (contact tracing isolation).

In network terminology, those intervention measures change the topological structure of the social-contact network consisting of nodes and links (Luo & Jin, 2020), in which the nodes mimic individuals and links stand for social contacts between individuals among the population. Under household quarantine in China, every person almost stayed at home which substantially reduces their contacts with outside, making the contact structure present strong clustering effects. According to Chinese Statistic Yearbook (Chinese Statistic Yearboo, 2019), there are ten household types in the household structure and the contact structure in the scenario of household quarantine (every individual only contacts with its family members) exhibit a unimodal degree distribution (see Fig. 1(a)), where the node degree is the number of contacts an individual has. On the other hand, contact tracing isolation mainly traces the individuals having contacted with individuals at incubation state. Upon the incubated individuals being confirmed, all individuals having contacted with them are isolated. In particular, the traced high-risk contacts from individuals at incubation and asymptomatic state to the susceptible are the key ones leading to the increase in infection.

As COVID-19 has aroused a great concern worldwide, there has been a sharp rise in the number of mathematical models to study its dynamics. In terms of the well-mixed models, Tang et al. assessed the transmission risk of COVID-19 spread at the early stage of epidemic by including the intervention measures (Tang et al., 2020a, 2020b). Yang et al. combined compartmental models with artificial intelligence (AI) approaches to predict the epidemics trend of COVID-19 in China (Yang et al., 2020). Some researchers analyzed the effects and effectiveness of quarantine and isolation on COVID-19 outbreak in Guangdong province China (Hu et al., 2020) and in China (Hou et al., 2020; Tang et al., 2020c). The role and effectiveness of social distance in containing COVID-19 was also discussed in some papers (Anderson et al., 2020; Leung, Wu, Liu, & Leung, 2020). In addition, in terms of network-based models, Xue et al. used a degree-based network model to study the transmission of COVID-19 in Wuhan, Toronto and Italy (Xue et al., 2020). Anyway, in the well-mixed models, it is assumed that the number of contacts are constant or man-made time-varying (Roda, Varughese, Han, & Li, 2020; Tang et al., 2020a), ignoring the real time-varying contact pattern between individuals. Although the degree-based network model considers the degree distribution—one of network topological characteristics, it is difficult for this model to characterize the clustering effects and high-risk contacts or depict two crucial intervention measures of household quarantine and contact tracing isolation. It is however remarkable that the pairwise model—one of the network-based models—could simultaneously including the evolution of individuals and contacts (Keeling, 1999), has a huge advantage of characterizing those two interventions taken in China over the well-mixed models (Li, 2018) and other network-based models, such as the degree-based models (Pastor-Satorras & Vespignani, 2001).

In this paper, we proposed a pairwise model, able to explicitly considering contact tracing isolation and household quarantine, to analyze the impact of NPIs on COVID-19 spread in China. As Fig. 2(A) indicated in (Ali et al., 2020), intervention measures on COVID-19 at the early stage in China changed with time until January 31, 2020. To fit a stable dynamic process for an ordinary differential equation, we chose the real data after January 31, 2020. We based on data of COVID-19 infections collected from National Health Commission of China (NHCC) (The National Health Commission of P. R. China) to identify the parameters of the proposed model and calculated the epidemic quantities: the reproduction number, peak time, epidemic size and epidemic duration (Luo & Jin, 2020). These quantities determine the potential and severity of an outbreak and are in favor of evaluation of control measures. Note that these quantities are based on the same control strategies at that time before the effects of imported cases begin to be dominant (the times of the first confirmed imported case outside and in Hubei province are March 18th and March 31st, respectively (The National Health Commission of P. R. China)). It is indicated that the

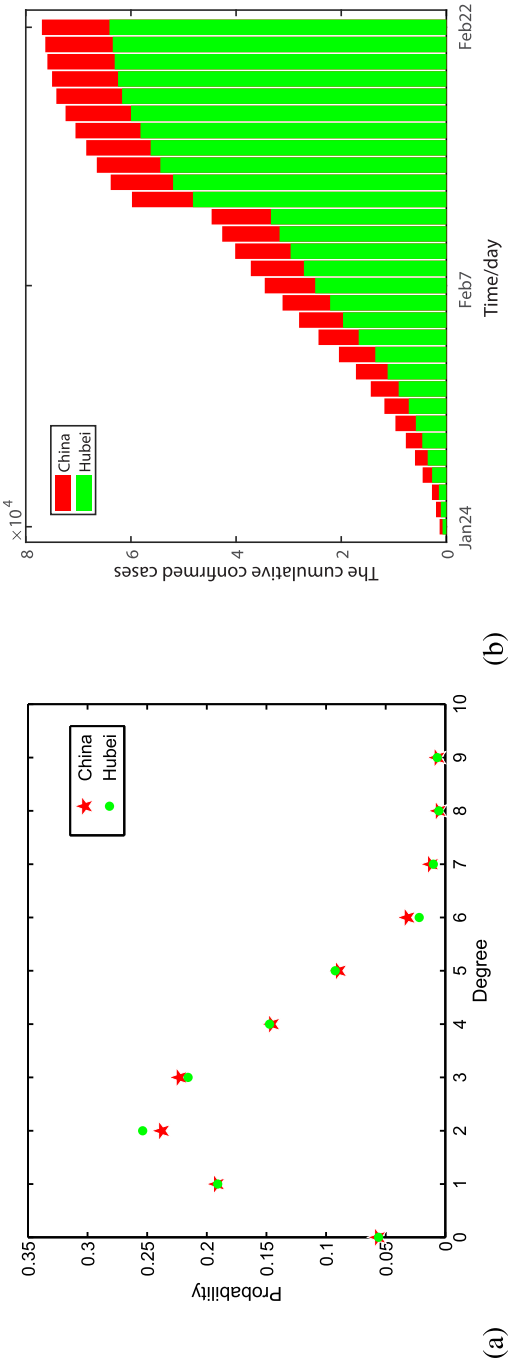

Fig. 1. (a)Degree distributions for China and Hubei province. (b) Real data of cumulative confirmed cases in China (red) and Hubei province (green).

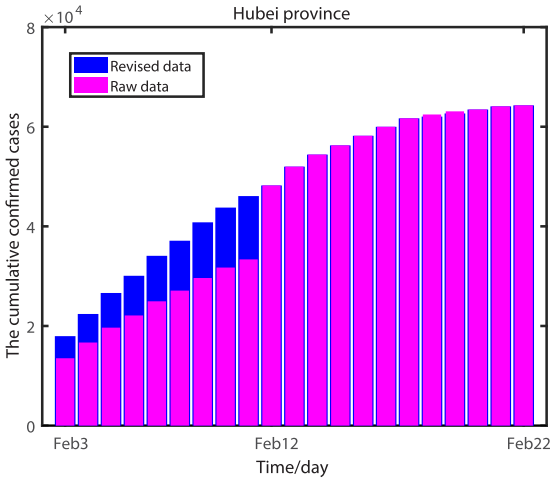

Fig. 2. Comparison between reported data and revised data of cumulative confirmed cases of COVID-19 in Hubei province.

traced high-risk contacts from incubated to susceptible decrease under NPIs, regardless of infected cases. The sensitivity analysis also showed that the enforcement of household quarantine measures decrease the reproduction number and epidemic prevalence. We also gave the resumption time of production via our model, which accords with the real time. Additionally, since Hubei province is the epicenter of COVID-19 outbreak and has suffered the most severe damage by the epidemic with lack of medical resources to quarantine exposed and suspect cases. As Fig. 1(b) shown, the reported cases of COVID-19 infections in Hubei province account for about 95% of the infections across China. In comparison, other regions outside Hubei possess sufficient medical resources and have a relatively minor burden of infected individuals due to imported cases from Hubei. Therefore, the two regions would be addressed separately in the following.

2. Materials and methods

2.1. Data collection and description

As Fig. 2(A) indicated in (Ali et al., 2020), intervention measures on COVID-19 changed with time until January 31, 2020 in China. In order to estimate relatively accurate parameters in an ordinary differential equation, we collected all the daily reported data on COVID-19 infections during the period between February 3rd–22nd, 2020 from NHCC, where all the laboratory confirmed cases, suspected cases, death-caused by COVID-19, cured cases and close contacts of population with confirmed infections are defined (The National Health Commission of P. R. China; The platform of 2019-nCov, 2019).

Since the diagnosed approaches changed on 12th February—about 12000 clinically diagnosed cases are counted in the cumulative confirmed cases reported on that day in Hubei province (Official news on the chan, 2020)—there exists a singular point for cumulative confirmed cases on February 12th, 2020 (see Fig. 1(a)). In order to obtain relatively reliable data, we allocated the sudden increment on 12th February to each day in the preceding week in proportion to the original daily increments of confirmed cases in these days (see Fig. 2 for the revised data).

The clustering coefficient (denoted by φ) of a network is defined as the ratio of the number of triangles to the number of triples in the network (Keeling, Rand, & Morris, 1997; Rand, 1999). In China, under the household quarantine measure for COVID-19 outbreak, citizens are suggested staying at their homes and could only be allowed to go outdoors for essential purposes such as buying necessary supplies and medicines. In this case, the underlying network structure is highly clustered as the individuals within each family are closely connected whereas there are relatively sparse connections among families. As a result, the stricter the household quarantine measure, the smaller the number of inter-family connections, and hence the larger the clustering coefficient of the network. That is, the clustering coefficient can be used to characterize the isolation intensity of household quarantine measure for COVID-19 outbreak. The more intensive the household quarantine measure, the larger the clustering coefficient. According to the family distribution in the China Statistic Yearbook (Chinese Statistic Yearboo, 2019), we obtained that if every family is completely isolated, then $\varphi = 0.5$; whereas if only one family member is allowed to go out for living necessities, then $\varphi = 0.42$ (see A for details), which is closest to the case of household quarantine in China at that time. Therefore, in this paper we used the value of $\varphi = 0.42$ to represent the clustering coefficient of the network during the period of COVID-19 outbreak in China.

Table 1
Model variables and their definitions.

variable	Definition
$[S]$	The number of unquarantined susceptible individuals
$[S_0]$	The number of quarantined susceptible individuals (i.e.,the number of susceptible individuals who have been contacted with confirmed individuals)
$[E]$	The number of unquarantined incubation individuals
$[E_0]$	The number of quarantined incubation individuals (i.e.,the number of incubation individuals who have been contacted with confirmed individuals)
$[A]$	The number of unquarantined asymptomatic individuals
$[A_0]$	The number of quarantined asymptomatic individuals
$[I]$	The number of confirmed individuals
$[R]$	The number of recovered individuals
$[SS]$	Twice the number of links between nodes with S state and S state
$[SE]$	The number of links between nodes with S state and E state
$[SA]$	The number of links between nodes with S state and A state
$[EE]$	Twice the number of links between nodes with E state and E state
$[EA]$	The number of links between nodes with E state and A state
$[AA]$	Twice the number of links between nodes with A state and A state
$[SSE]$	The number of triples with the joint structure $S-S-E$
$[SSA]$	The number of triples with the joint structure $S-S-A$
$[ASE]$	The number of triples with the joint structure $A-S-E$
$[ESA]$	The number of triples with the joint structure $E-S-A$ and $[ASE]=[ESA]$
$[ESE]$	Twice the number of triples with the joint structure $E-S-E$
$[ASA]$	Twice the number of triples with the joint structure $A-S-A$
$[EEE]$	Twice the number of triples with joint structures $E-E-E$
$[SEE]$	The number of triples with the joint structure $S-E-E$
$[EEA]$	The number of triples with the joint structure $E-E-A$
$[AAE]$	The number of triples with the joint structure $A-A-E$
$[SAE]$	The number of triples with the joint structure $S-A-E$

2.2. Pairwise epidemic model for COVID-19 infections

The main purpose of this article is to investigate how the novel coronavirus spreads across China after the Chinese government on January 23rd, 2020 made the level-1 public health emergency response to fight against the COVID-19 epidemic by a series of all-out efforts, such as sealing off the Wuhan city, restricting travels, contact tracing, household isolation, and putting those who had close contacts with infected ones in quarantine for medical observation. The measures of household isolation, contact tracing and targeted quarantine often result in relatively intensive family clusters and contact reduction between susceptible individuals and confirmed cases of the disease. Traditional compartment epidemic models typically fail to account for the cluster structure and contact changes among individuals in the population. In this situation, one of the alternatives can be a pairwise epidemic model, where the whole population is considered as a contact network with each node representing an individual of the population and each edge (or link) standing for a contact between a pair of individuals (see Table 1).

Here we proposed a pairwise network epidemic model to incorporate the effects of family clusters and quarantine-caused contact cutting on the disease transmission. As demonstrated in Fig. 3, we classified all of the individuals (nodes) into eight different states: unquarantined susceptible (S), quarantined susceptible (S_0), unquarantined incubation (E), quarantined incubation (E_0), unquarantined asymptomatic (A), quarantined asymptomatic (A_0), confirmed infected (I) and removed (R). Here, the removed compartment contains disease-induced death and recovered. We assumed that individuals in either state of E, A and I are infectious; however, since infected individuals will be immediately put in quarantine once they are confirmed, the individuals in state I are not allowed to transmit the disease. The transitions between different states are as follows. Each S individual turns into the E state when it is infected at the transmission rate β_1 through a contact with an E individual, while at rate β_2 by an A individual. After a incubation period of $1/q$ days, every E individual jumps into either the A class with probability $1-p$ or the I class with probability p. The individuals in states I and A run into the R class at the rate γ_1 and γ_2,

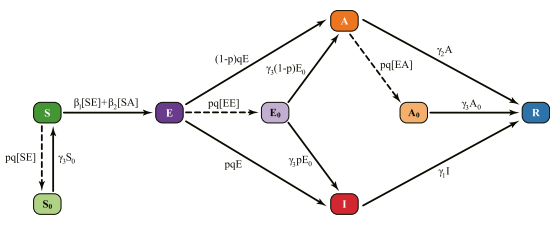

Fig. 3. The flow diagram of pairwise model for COVID-19 epidemic.

respectively. Meanwhile, the individuals in states S, E and A will be isolated and turn into the S_0, E_0 and A_0 states, respectively, once any one of their neighbors is confirmed to have been infected. Each individual of state S_0 or A_0 will be released from quarantine and enter into the S or R state, respectively, after an observation period of $1/\gamma_3$ days. Whereas each individual in state E_0 will be released from quarantine after an observation period of $1/\gamma_3$ days, entering into either the A class with probability $1-p$ or the I class with probability p. It is worth noting that the close contacts among individuals of different states are denoted by the pairwise variables in square brackets as shown in Fig. 3. Parameters used in the model have been summarized in Table 2.

Using the notations defined in Table 1, the pairwise model for COVID-19 epidemic can be described as follows. Firstly, the time evolution of the number of individuals in different classes is given by the following equations.

$$\frac{d[S](t)}{dt} = \gamma_3[S_0] - \beta_1[SE] - \beta_2[SA] - pq[SE], \tag{1}$$

$$\frac{d[E](t)}{dt} = \beta_1[SE] + \beta_2[SA] - q[E] - pq[EE], \tag{2}$$

$$\frac{d[I](t)}{dt} = pq[E] + p\gamma_3[E_0] - \gamma_1[I], \tag{3}$$

$$\frac{d[A](t)}{dt} = (1-p)q[E] + (1-p)\gamma_3[E_0] - \gamma_2[A] - pq[EA], \tag{4}$$

$$\frac{d[R](t)}{dt} = \gamma_1[I] + \gamma_2[A] + \gamma_3[A_0], \tag{5}$$

$$\frac{d[S_0](t)}{dt} = pq[SE] - \gamma_3[S_0], \tag{6}$$

$$\frac{d[E_0](t)}{dt} = pq[EE] - \gamma_3[E_0], \tag{7}$$

$$\frac{d[A_0](t)}{dt} = pq[EA] - \gamma_3[A_0]. \tag{8}$$

Since the time evolution of the number of individuals in each class given by Eqs. (1)–(8) have included the numbers of different types of links, in order to close Eqs. (1)–(8) we also presented in Eqs. 9–14 the time evolution of the number of links of each type.

$$\frac{d[SS]}{dt} = -2\beta_1[SSE] - 2\beta_2[SSA] - 2pq[SSE], \tag{9}$$

Table 2
Definition of parameters and their values (unit time: day).

Parameter	Definition	China (except Hubei)	95% CI	Hubei	95% CI	Data source
β_1	transmission rate by E individuals	0.197	[0.194, 0.201]	0.207	[0.205, 0.209]	MCMC
β_2	transmission rate by A individuals	0.111	[0.109, 0.113]	0.201	[0.198, 0.204]	MCMC
p	probability of showing symptoms	0.688	[0.686, 0.701]	0.667	[0.664, 0.670]	MCMC
$1/q$	incubation period for E individuals	5	–	6	–	Tang et al. (2020b)
γ_1	The transferred rate of individuals from I to R	1/10	–	1/10	–	[21, 22]
γ_2	The transferred rate of individuals from A to R	1/12	–	1/12	–	[21, 22]
$1/\gamma_3$	quarantine period for S, E and A individuals	14	–	14	–	Tang et al. (2020b)
N	total population	1.33621e9	–	5.9170e7	–	Chinese Statistic Yearboo (2019)
φ	clustering coefficient	0.42	–	0.42	–	Chinese Statistic Yearboo (2019)
n	average node degree	3.1	–	3.04	–	Chinese Statistic Yearboo (2019)

$$\frac{d[SE]}{dt} = \beta_1[SSE] + \beta_2[SSA] - \beta_1[SE] - q[SE] - \beta_1[ESE] - \beta_2[ASE] - pq[SEE] - pq[ESE], \tag{10}$$

$$\frac{d[SA]}{dt} = (1-p)q[SE] - \gamma_2[SA] - \beta_2[SA] - \beta_1[ESA] - \beta_2[ASA] - pq[ESA] - pq[SAE], \tag{11}$$

$$\frac{d[EE]}{dt} = 2\beta_1[SE] + 2\beta_1[ESE] + 2\beta_2[ASE] - 2q[EE] - 2pq[EEE], \tag{12}$$

$$\frac{d[EA]}{dt} = (1-p)q[EE] + \beta_2[SA] + \beta_1[ESA] + \beta_2[ASA] - q[EA] - \gamma_2[EA] - pq[EEA] - pq[EAE], \tag{13}$$

$$\frac{d[AA]}{dt} = 2(1-p)q[EA] - 2\gamma_2[AA] - 2pq[AAE]. \tag{14}$$

For detailed explanation of the biological meaning of Eqs. 1–14, the readers are referred to B. The appearance of the numbers of triples of different types suggests one to close the model by utilizing the moment closure approximation formula by Keeling et al. (House & Keeling, 2011; Keeling et al., 1997) as follows:

$$[XYZ] = \frac{n-1}{n}\frac{[XY][YZ]}{[Y]}\left(1-\varphi+\varphi\frac{N}{n}\frac{[XZ]}{[X][Z]}\right), \qquad X,Y,Z \in \{S,E,A\}, \tag{15}$$

where n is the average degree, φ is the clustering coefficient, and N is the number of nodes in the underlying network.

2.3. The reproduction number

The reproduction number R, which is defined as an average numbers of secondary cases that one case produces during the course of its infectious period, determines the increase growth rate of an emerging infectious disease. Sometime it can be used to judge whether or not one disease breaks out. Adopting the method proposed by Yang and Xu (Yang & Xu, 2019) or the next generation matrix method (Van den Driessche & Watmough, 2002), the expression R takes the form of

$$R = (1-\varphi)(n-1)\left(\frac{\beta_1}{\beta_1+q} + \frac{(1-p)q\beta_2}{(\beta_1+q)(\beta_2+\gamma_2)}\right) = R^{[SE]} + R^{[SA]}, \tag{16}$$

where

$$R^{[SE]} = (1-\varphi)(n-1)\frac{\beta_1}{\beta_1+q}, \tag{17}$$

and

$$R^{[SA]} = (1-\varphi)(n-1)\frac{(1-p)q\beta_2}{(\beta_1+q)(\beta_2+\gamma_2)}. \tag{18}$$

Since the pairwise epidemic model (1-14) is a network model, we mainly concerned the variants of the infected links. The biological meanings of R can be regards as a sum of two quantities. $R^{[SE]}$ is the number of secondary infected links that one exposed link will generate in an entirely susceptible link $(1-\varphi)(n-1)$ during its lifespan $\frac{1}{\beta_1+q}$ as exposed. Note that, $\frac{(1-p)q}{\beta_1+q}$ denotes a probability that one exposed link survives the expose link and progresses into the infected link. Hence, $R^{[SA]}$ is the number of secondary infected links that one infected link will produce in an totally susceptible links during its lifespan $\frac{1}{\beta_2+\gamma_2}$.

2.4. MCMC estimation

Here, we briefly introduced the adaptive Metropolis-Hastings (M-H) algorithm of Markov Chain Monte Carlo (MCMC) simulations (Gamerman & Lopes, 2006), by which we estimated parameters and fit data to our model (1-14).

Since we used the reported cumulative confirmed cases $z(T)$ for data fitting, we denoted according to Eq. (3) the expectation of the cumulative confirmed individuals by

$$\bar{z}(T) = \sum_{t=1}^{T}(pq[E](t) + p\gamma_3[E_0](t)), \quad T = 1, 2, \ldots, 10, \tag{19}$$

and its variation with a small perturbation by

$$z(T) = \bar{z}(T) + \delta, \quad T = 1, 2, \ldots, 10, \tag{20}$$

where $z(1) = \bar{z}(1)$, $T = 1, 2, \ldots, 10$ corresponds to the days during the period February 3rd–10th, 2020, and δ is supposed to follow the normal distribution $N(0, \epsilon^2)$ with $\epsilon = 1000$. Given the parameter vector Θ, the maximum likelihood function for the revised data $Z = \{z(T)\}_{T=0}^{10}$ is

$$L(Z|\Theta) = \Pi_{T=1}^{12}\frac{1}{\sqrt{2\pi}\epsilon}exp\left(-\frac{[z(T) - \bar{z}(T)]^2}{\epsilon^2}\right) \tag{21}$$

Then, the joint posterior distribution of the parameters is given by $P(\Theta|Z) \propto L(Z|\Theta)P(\Theta)$. Under the assumption of non-information prior distribution (Gamerman & Lopes, 2006), the joint posterior distribution of the parameters satisfies $P(\Theta|Z) \propto L(Z|\Theta)$. Furthermore, at each iteration step the new parameter vector Θ is updated through a random walk process and will be accepted with probability

$$\min\left\{1, \frac{P(\bar{\Theta}|Z)}{P(\Theta|Z)}\right\} \tag{22}$$

2.5. Parameters and initial conditions

The population size of China (except Hubei) and Hubei province is, respectively, about 1, 336, 210, 000 and 59, 170, 000 (Chinese Statistic Yearboo, 2019), so we set $[S](t_0)$ to be 1, 336, 210, 000 and 59, 170, 000 for China (except Hubei) and Hubei province, respectively. Here, t_0 stands for the starting date February 3rd, 2020, when the number of the cumulative confirmed cases in Hubei and China (except Hubei) is 13, 522 and 6, 028, respectively, and the number of the cumulative recovered cases in Hubei and China (except Hubei) is 810 and 169, respectively. Thus, we set $[I](t_0)$ to 12, 712 (= 13, 552−810) and 5, 837 (= 6, 028−169) for Hubei and China (except Hubei), respectively. Similarly, we determined the initial values of $[S_0]$, $[E_0]$ and $[A_0]$ according to the reported data of suspected cases (identified by contact tracing) who are quarantined for medical observation. The quarantined individuals are isolated for 14 days, so we set $\gamma_3 = 1/14$. According to literature (Tang et al., 2020b), γ_2 is about 1/12, and the incubation period is about 6 days. However, as the test technique and supply of medical resources improve, especially outside Hubei province, we set the incubation period $1/q$ to 5 days for China (except Hubei) and 6 days for Hubei. Moreover, based on the surveillance data we calculated the average of hospitalization period for confirmed cases to be about 10 days. Thus we set $\gamma_1 = 1/10$. The average node degree n in the network is the average number of contacts an individual has in the population. According to the degree distribution (Fig. 1(a)) obtained from the data published on the Chinese Statistic Yearbook (Chinese Statistic Yearboo, 2019), we got $n = 3.1$ for China (except Hubei) and $n = 3.04$ for Hubei. We determined the initial numbers of *EE*-links, *EA*-links and *AA*-links based on the following approximation formula,

$$[XY] = n \times [X] \times \frac{[Y]}{N}, \quad X,\ Y \in \{S, E, I, A, R\}, \tag{23}$$

where $[X]$ or $[Y]$ represents the number of nodes in state X or Y and $[XY]$ represents the number of XY-links. Note that we have set the initial numbers of *EE*-links, *EA*-links and *AA*-links to 1 since formula (23) gives a real number less than 1.

All the parameters and initial values are summarized in Tables 2 and 3.

3. Results

3.1. Fitting results

Based on the analysis of the diagnosed cases of COVID-19 infection reported in (Health Commission of Hubei Province), we adopted the MCMC method (Gamerman & Lopes, 2006) for 10000 iterations with a burn-in of 6000 iterations to fit the model (1-14) and estimated the parameters and the initial conditions of variables (see Tables 2 and 3). The Metropolis-Hasting (MH) algorithm is used to adapt and readjust the MCMC procedure. To assess the goodness-of-fit for the validation data, we used the first half of the data (for dates from February 3rd to 12th in 2020) as training data and the second half (for dates from February 13th to 22nd in 2020) as validation data. The result presented in Fig. 4 shows a good fitting between the model

Table 3
Initial values used in the model.

Parameter	Definition	China (except Hubei)	95% CI	Hubei	95% CI	Data source
$[S](t_0)$	initial number (#) of S individuals	N	–	N	–	Chinese Statistic Yearboo (2019)
$[E](t_0)$	initial # of E individuals	6366	[6,126, 6546]	28,636	[27,992, 29,280]	MCMC
$[I](t_0)$	initial # of I individuals	5837	–	12,712	–	[21, 22]
$[A](t_0)$	initial # of A individuals	1178	[1,098, 1258]	8138	[7,980, 8296]	MCMC
$[S_0](t_0)$	initial # of S_0 individuals	103,077	–	57,731	–	[21, 22]
$[E_0](t_0)$	initial # of E_0 individuals	726	–	407	–	[21, 22]
$[A_0](t_0)$	initial # of A_0 individuals	726	–	407	–	[21, 22]
$[SE](t_0)$	initial # of SE-links	2446	[2,396, 2501]	24,814	[24,188, 25,439]	MCMC
$[SA](t_0)$	initial # of SA-links	2105	[1,986, 2224]	27,467	[26,743, 28,191]	MCMC
$[SS](t_0)$	initial # of SS-links	1,476,689,583	[1,437,605,190, 1,515,773,970]	37,709,952	[35,087,662, 40,332,242]	MCMC
$[EE](t_0)$	initial # of EE-links	1	–	1	–	(23)
$[EA](t_0)$	initial # of EA-links	1	–	1	–	(23)
$[AA](t_0)$	initial # of AA-links	1	–	1	–	(23)

solution and real data, well suggesting the epidemic trend in China under the intervention of household quarantine containment strategies. According to the estimated parameter values and initial conditions as given in Tables 2 and 3, we obtained the mean value of the reproduction number $R = 1.345$ with 95% CI of [1.230, 1.460] for the Hubei province and $R = 1.217$ with 95% CI of [1.207, 1.227] for the regions outside Hubei province. Note that they are actually the effective reproduction number. Since the calculation of the basic reproduction number is based on the initial conditions at the start of epidemic, the reproduction number we calculated took the February 3rd as a starting point, which represents the scenario on February 3rd when the strict control measures have been implemented. It may be one of the reasons that they were smaller than those published in some literature (Shao & Shan, 2020; Shen, Peng, Xiao, & Zhang, 2020; Tang et al., 2020b; You et al., 2020), as those values of estimated parameters in the model heavily depend on the choice of the initial surveillance data. The other reason may be that most of existing results ignore clustering effects induced by home quarantine as well as dynamical correlations between susceptible individuals and infectious individuals. Moreover, we readjusted the above processes and got good fitting results in the top four provinces in China: Guangdong, Henan, Hunan, and Jiangxi, respectively (see Fig. 11 in C). In what follows, we mainly employed our model with estimated parameters to study the epidemic behavior of COVID-19 infections in China (except Hubei) and in Hubei.

3.2. The estimation of epidemiological quantities

Note that as Section 1 stressed, the epidemiological quantities estimated in this paper are based on the control strategies at that time before the effects of imported cases begin to be dominant (the times of the first reported confirmed imported case outside and in Hubei province are March 18th and March 31st, respectively (The National Health Commission of P. R. China)). If not, they will change accordingly as long as the disease control measures are modified. Applying the estimated parameter values, we obtained that the outbreak of COVID-19 in China (except Hubei province) peaks on February 7th, 2020 (95% CI between February 3rd and February 11th) (Fig. 5(a)), whereas the peak time of COVID-19 infections in Hubei province is on February 13th, 2020 (95% CI between February 9th and February 17th) (Fig. 5(b)). The actual peak times are on February 10th, 2020 and on February 17th, 2020 in those two zeros, respectively (The National Health Commission of P. R. China; The platform of 2019-nCov, 2019). The estimated errors are less then 5 days. The outbreaks of COVID-19 are estimated to last until March 22nd, 2020 (95% CI between March 17th and March 27th) in China (except Hubei province) and until April 7th, 2020 (95% CI between March 30th and April 15th) in Hubei province before the effects of imported cases begin to be dominant. The corresponding epidemic sizes at that time are estimated as 13,359, with 95% CI of [12, 678, 14, 752] in China (except Hubei) and 70,160 with 95% CI of [69, 351, 71, 755] in Hubei province. The real epidemic sizes in two regions are 13,293 and 64,142, respectively, and the relative errors of the epidemic sizes are 0.005 and 0.094, less then 1%. The primary reason for the COVID-19 epidemic ending late with a larger epidemic size in Hubei province is that the ongoing exposure of human population to COVID-19 contagion in Hubei province is relatively higher than other regions in China. Moreover, limited resources of medical and health cares in Hubei province, especially in Wuhan, the epicenter of the epidemic, have significantly increased the risk of COVID-19 propagation. It could be easily seen that the crucial epidemiological quantities given by our model have good agreements with the real data.

In addition, Fig. 6 shows the evolution of the high-risk contacts from infectious to susceptible $[SE](t) + [SA](t)$ over time under NPIs in China except Hubei province and Hubei province since February 2nd, 2020. It could be seen that those high-risk contacts decrease under NPIs, although infections still increase from February 2nd, 2020 to the peak time (see Fig. 5). In fact, the crucial factors determining the prevalent results of infectious diseases, are high-risk contacts from infectious to susceptible rather than the number of infectious individuals. Fig. 6 verified the effectiveness of NPIs implemented in China.

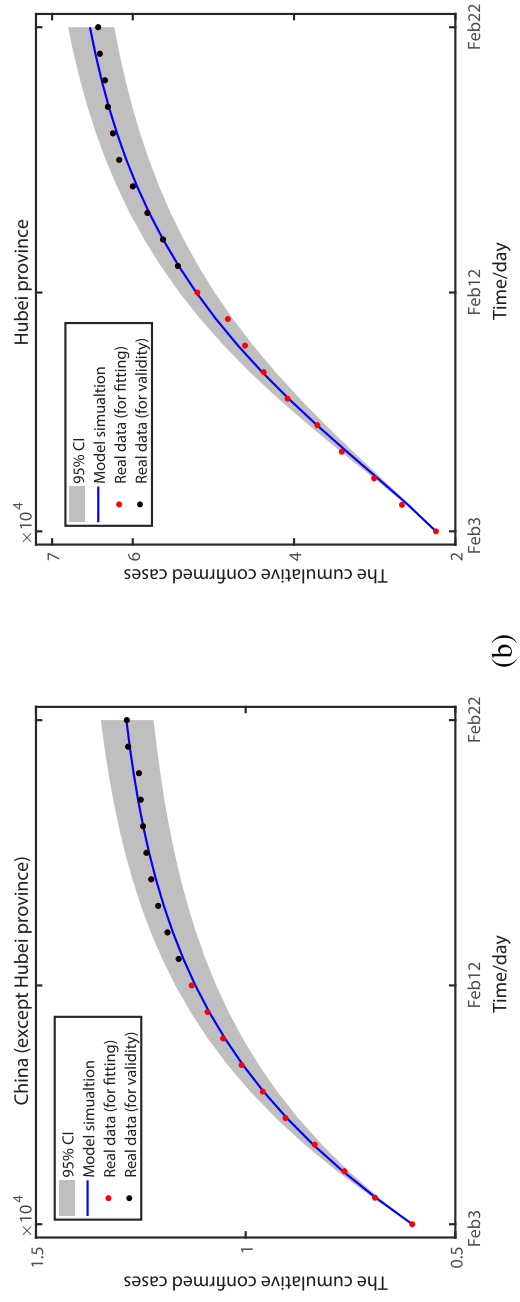

Fig. 4. The fitting results of our model to real data of COVID-19 infections in China exclusive of Hubei province (a) and that in Hubei province (b). The grey area marks the 95% CI of MCMC estimations. The red points are training data for parameter fitting, while the black points are real data for validation.

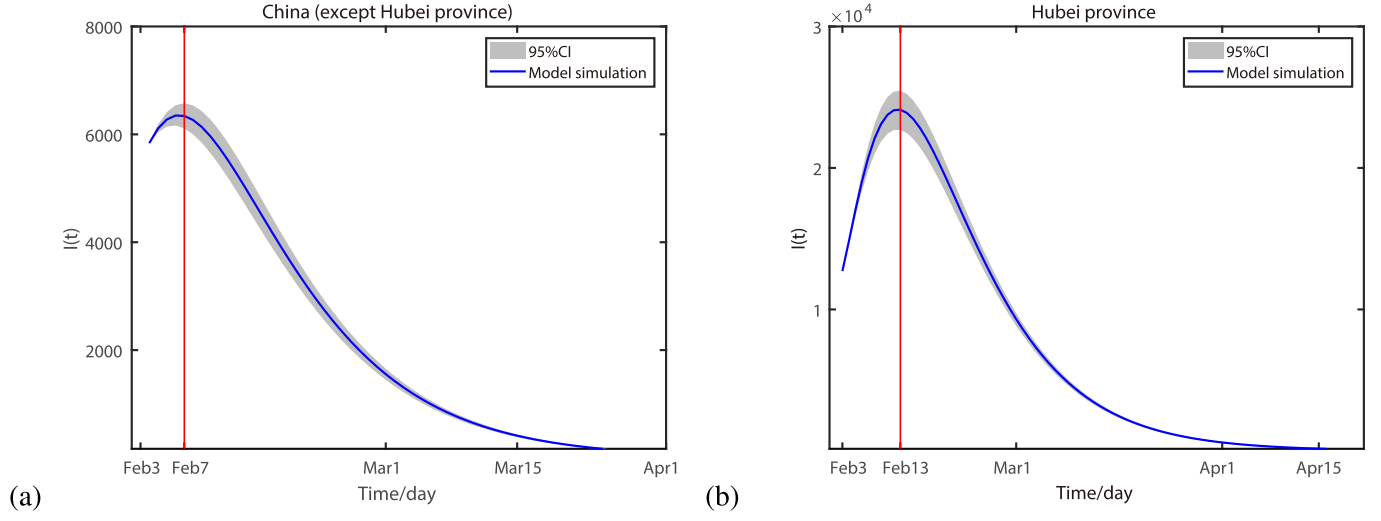

Fig. 5. The time series of the number $I(t)$ of confirmed individuals in China except Hubei province (a) and Hubei province (b). The red vertical lines in the plots mark the peak time for the COVID-19 spread of which the trajectory is plotted in blue curves.

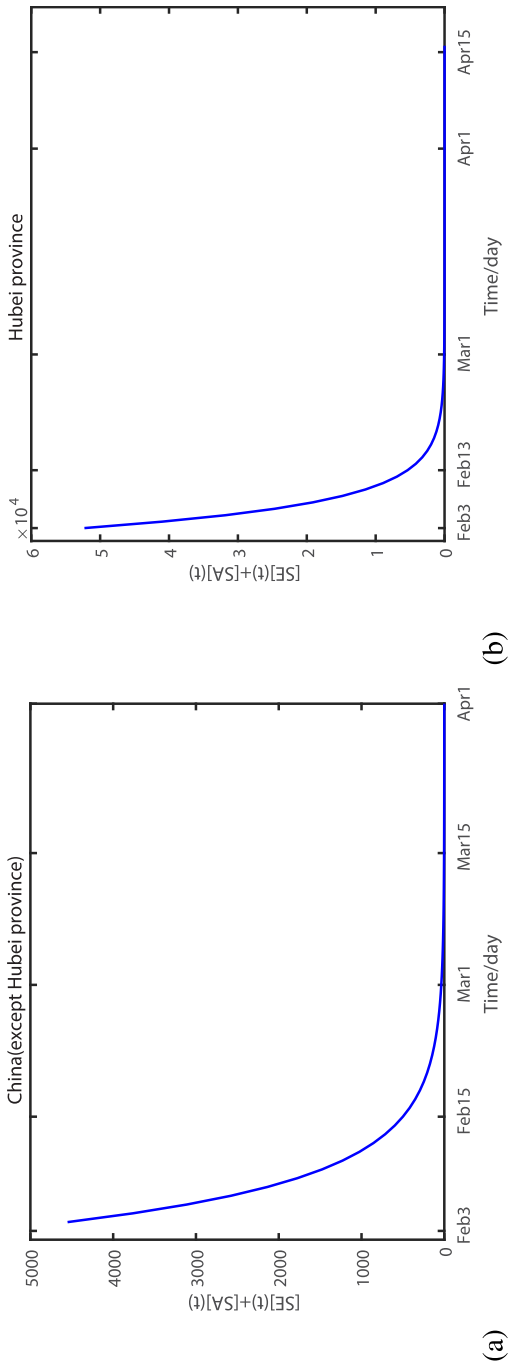

Fig. 6. The time series of $[SE](t) + [SA](t)$ in China except Hubei province (a) and Hubei province (b).

3.3. The sensitive analysis

In the face of an emerging disease, the Chinese government has taken many efficacious strategies to control the disease transmission, such as traffic restrictions, isolations, home quarantine, increasing medical resource and propaganda education, etc. To examine and evaluate the potential efficacy of these strategies, we did the sensitivity analysis for two vital model parameters q and φ, which reflect the intensity of detection and isolation, respectively.

Fig. 7 and Tables 4 and 5 show that shortening the incubated period $1/q$ from 7 days to 3 days decreases the epidemic size and advances the epidemic duration by about one week. In addition, the peak arrival time also advances. As the advance time is little, we did the partial rank correlation analysis of the peak arrival time on the incubation period. We choose 1000 samples from 7 days to 3 days of the incubation period $1/q$ with other parameters following the normal distribution in Table 2. It is found that the difference of the advance time is not significant in China (except for Hubei province) but significant in Hubei province with PRCCs $= 0.993$ and $p = 0.0234$. It means that shortening the incubation period actually advances the peak arrival time in Hubei province. Overall, increasing the detection rates as an increase of q help timely trace and isolate individuals having close contacts with confirmed cases, which is in favor of mitigating COVID-19 transmission.

As Section 2.1 introduced, the intensity of isolation under the household quarantine measure is characterized by the clustering coefficient φ. Tables 4 and 5 show that, with the increase of φ or the intensity of isolation in China, the epidemic size decrease, the peak time advances and the epidemic duration is shortened. It is worth mentioning that when the intensity of isolation increases, the exposure proportion of susceptible individuals to COVID-19 decreases. It means that both clustering coefficients φ and the exposure proportion of susceptible individuals may affect the spread of COVID-19. Fig. 8 presents the contour figure of the epidemic size with respect to φ and the exposure proportion of susceptibles. We could see that with the decrease of φ from 0.5 to 0.1 or the increase of the susceptible exposure proportion from 0.35 to 0.55, the epidemic sizes increase accordingly. In next section, we will further discuss their effects on COVID-19 combined with control measures.

4. Discussion

COVID-19 was rife in China. The Chinese government has taken many extreme strategies after January 23rd, 2020, which resulting in strong correlation between incubated individuals, asymptomatic individuals, confirmed cases and susceptible people. Moveover, Chinese government set up a series of tracking system in order to follow the trail of each confirmed case and found close contacts of population. Once they found these close contacts of ones, they would be quarantined somewhere (at home or a central isolation). In consideration of the above features, we built a COVID-19 infection model by pairwise approach which couples with the Chinese family network structures and clustering effects, and considered the instant cut-off of close contacts with confirmed cases.

During the prevalence of COVID-19 in China, what people concern was when they could return to work. As shown in Section 2.1, the clustering coefficient, as one of the network characteristic quantities, reflects the gathering degree of individuals and plays a key role in the modelling process. In our study, as described above, $\varphi = 0.42$ represents the gathering degree in the current household quarantine measure. Once returning to work, people interconnects closely and the clustering coefficient decreases (see A). Meanwhile, Fig. 8 indicates isolation measure may change φ and the exposure proportion of susceptible individuals. Fig. 9 provides the different results with relaxing isolation measure at different times based on our network model. We have seen from Fig. 9(a) that in terms of the total population ($[S](0) = N$) in China except Hubei province, when increasing φ from 0.42 to 0.5 or decreasing φ from 0.42 to 0.1 on the February 7th, the epidemic size and the epidemic duration have small changes. On the contrary, in the terms of partial population ($[S](0) = 0.5N$), when decreasing φ from 0.42 to 0.1 on February 7th, the epidemic size increases from 13,399 to 18,445 and the epidemic duration delays from 53 days to 79 days, and when the same change of φ appears on March 1st, the epidemic size and epidemic duration changes slightly (see Fig. 9 (b)). It suggests that for areas outside Hubei, it would be better to maintain the current control measures and not to resume work and production until March 10th—after two incubation periods. The suggested time is consistent with the notification time for resumption of work and production, March 6th, when National Health Commission of China announced (The National Health Commission of P. R. China Chan & Yuen, 2020). Similarly, in Hubei province, when increasing φ from 0.42 to 0.5 or decreasing φ from 0.42 to 0.1 on the February 13th in terms of total population (see Fig. 9 (c)), the epidemic size and the epidemic duration have trivial change, whereas when taking the same changes in terms of partial population (Fig. 9 (d)), the epidemic size increases from 70,160 to 82,668 and the epidemic duration delays from 72 to 82 days. As the severity of COVID-19 in Hubei province, it may be plausible to resume work, production and other gathering activities in Hubei province until the end of April when the epidemic was basically brought to an end. It also corresponds to the notification time for resumption of work and production, April 24th, when Health Commission of Hubei Province announced (The National Health Commission of P. R. China).

The q, one of our model parameters, could reflect the strength of testing efforts and timeless of contact tracing isolation (Ali et al., 2020). Fig. 10(a) indicates that when intensifying the medical detection (q from 1/7 to 1/3) on February 7th, 2020 in China (except Hubei), the epidemic size decreased from 13,506 to 13,242 and the epidemic duration advances about one week. However, when taking the same detection change on 1st March and, the effects are small. Likewise, similar impacts happen in Hubei province. It means that the earlier the medical detection is promoted, the more timely close tracing isolation would be, the earlier COVID-19 would end and the smaller the epidemic size would be.

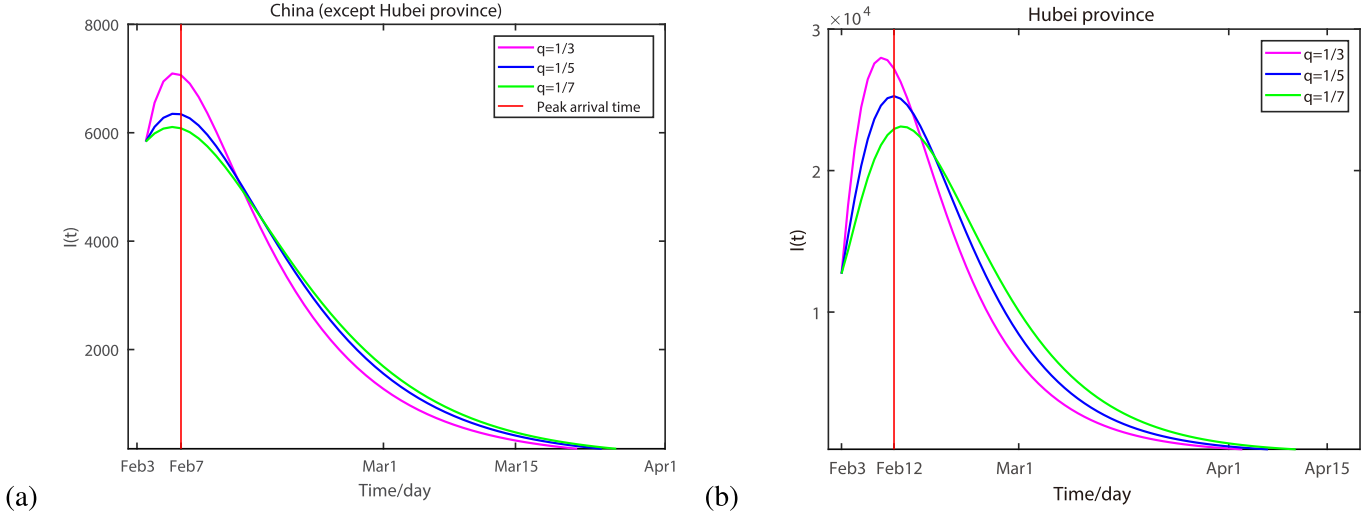

Fig. 7. The numbers of confirmed infected individuals change with different incubation periods $1/q$ for: (a) China (except Hubei); (b) Hubei province. The red vertical lines in the plots mark the peak time for the COVID-19 spread of which the trajectory is plotted in blue curves.

Table 4
Epidemic quantities for China except Hubei. The epidemic duration is counted starting from February 3rd, 2020.

Parameters	Peak time	Epidemic size	Epidemic duration
$q = 1/3$	2020.02.06	13,019	50 days
$q = 1/5$	2020.02.07	13,399	53 days
$q = 1/7$	2020.02.08	13,547	56 days
$\varphi = 0.1$	2020.02.07	13,864	55 days
$\varphi = 0.3$	2020.02.07	13,557	53 days
$\varphi = 0.5$	2020.02.07	13,301	52 days

Table 5
Epidemic quantities for Hubei province. The epidemic duration is counted starting from February 3rd, 2020.

Parameters	Peak time	Epidemic size	Epidemic duration
q = 1/3	2020.02.10	66,178	69 days
q = 1/5	2020.02.12	69,059	73 days
q = 1/7	2020.02.13	71,079	77 days
$\varphi = 0.1$	2020.02.14	71,498	77 days
$\varphi = 0.3$	2020.02.13	70,759	76 days
$\varphi = 0.5$	2020.02.13	69,755	75 days

As is known, a social-contact network consists of individuals and contacts between individuals in the population (Newman, 2010). When an infectious disease spreads in the social-contact network, the contact structure plays an vital role. In the well-mixed models, it is assumed that the contact structure is homogeneous, not able to describe the impact of heterogenous contacts between individuals on the spread of infectious diseases (Hou et al., 2020; Hu et al., 2020; Tang et al., 2020a, 2020b, 2020c; Yang et al., 2020). In contrast, the pairwise models take nodes with different status and contacts between nodes as the variables, considering evolution of individuals and relationships between individuals—especially clustering—with time (Keeling, 1999). Obviously, the pairwise model could explicitly characterize the close contacts between individuals and household quarantine by family clusters, and analyze their effects on COVID-19 spread. Therefore, it may be a relatively reasonable model to study COVID-19 in China. Even so, the theoretical analysis on this type of models is relatively difficult with coupling contact structure and epidemic dynamics.

Research has demonstrated that face mask could effectively decrease infected risk of COVID-19 as coarse respiratory droplets is one of the primary transmission routes of SARS-CoV-2 (Eikenberry et al., 2020). In other words, face mask could greatly reduce the transmission rate of COVID-19. In fact, they combined with other NPI have made a big contribution to containing COVID-19 in China and other countries (Chan & Yuen, 2020; Wang, Ng, & Brook, 2020). In our pairwise model, the edge variables describe the contact relationships between individuals with different state, and the effect of face mask could be characterized by the transmission rate. Given no available data of validity of face mask, research on their effects was not our concern, but could be a focus in the future work.

5. Conclusion

An emerging outbreak of COVID-19 has been effectively brought in control under NPIs in China, but still wreaks havoc in the world. Facing to the sudden epidemics, the Chinese government instantly has set up a series of extreme control measures including the closure of city, traffic restrictions, close tracing, holiday extensions and even household (family) quarantine since January 23rd, 2020. With the change of these policies, especially the closure of city and close contact tracing followed by household quarantine, the clustering phenomena occasionally exhibited in form of family clusters and hence the transmission patterns of COVID-19 infections had radically altered. Considering such signatures, we proposed a pairwise epidemic model on a network with relevant population structure to analyze the effects of NPIs on the dynamics of COVID-19 propagation in China. Coupling surveillance data started on February 3rd with our proposed model, the effective reproduction numbers were evaluated as 1.345 (95% CI [1.230, 1.460]) for Hubei province and 1.217 (95% CI [1.207, 1.227]) for China except Hubei. Obviously, they were smaller than those published in some literature (Shao & Shan, 2020; Shen et al., 2020; Tang et al., 2020b; You et al., 2020), as the control measures were implemented. It also gave further evidence that nonpharmaceutical interventions in China was effective and efficient. Interestingly, our model successfully mimics the time surveillance data from February 3rd to February 12th and the solution of model in China (see Fig. 4). Based our model, the estimated peak arrival time and cumulative confirmed cases are close to the real data of February 12th and the real number of cases of 72, 436. Sensitive analysis also indicated that household quarantine (characterized by φ) and timeless of close contact tracing isolation (characterized by q) make a great contribution to control of COVID-19 in China. Moreover, our results could be extended to do analysis for Guangdong, Hunan, Henan, and Jiangxi, as seen in Fig. 11. The COVID-19 infections in China has been effectively control under continuous timely and effective public health policies and containment measures. Our study has both significant theoretic and practical values for providing potential suggests to control COVID-19 for the other parts of world.

Fig. 8. Contour plots of the epidemic size with regard to the clustering coefficient φ and the initial exposure proportion of susceptibles $[S](t_0)/N$ in the whole population for (a) China except Hubei province and (b) Hubei province.

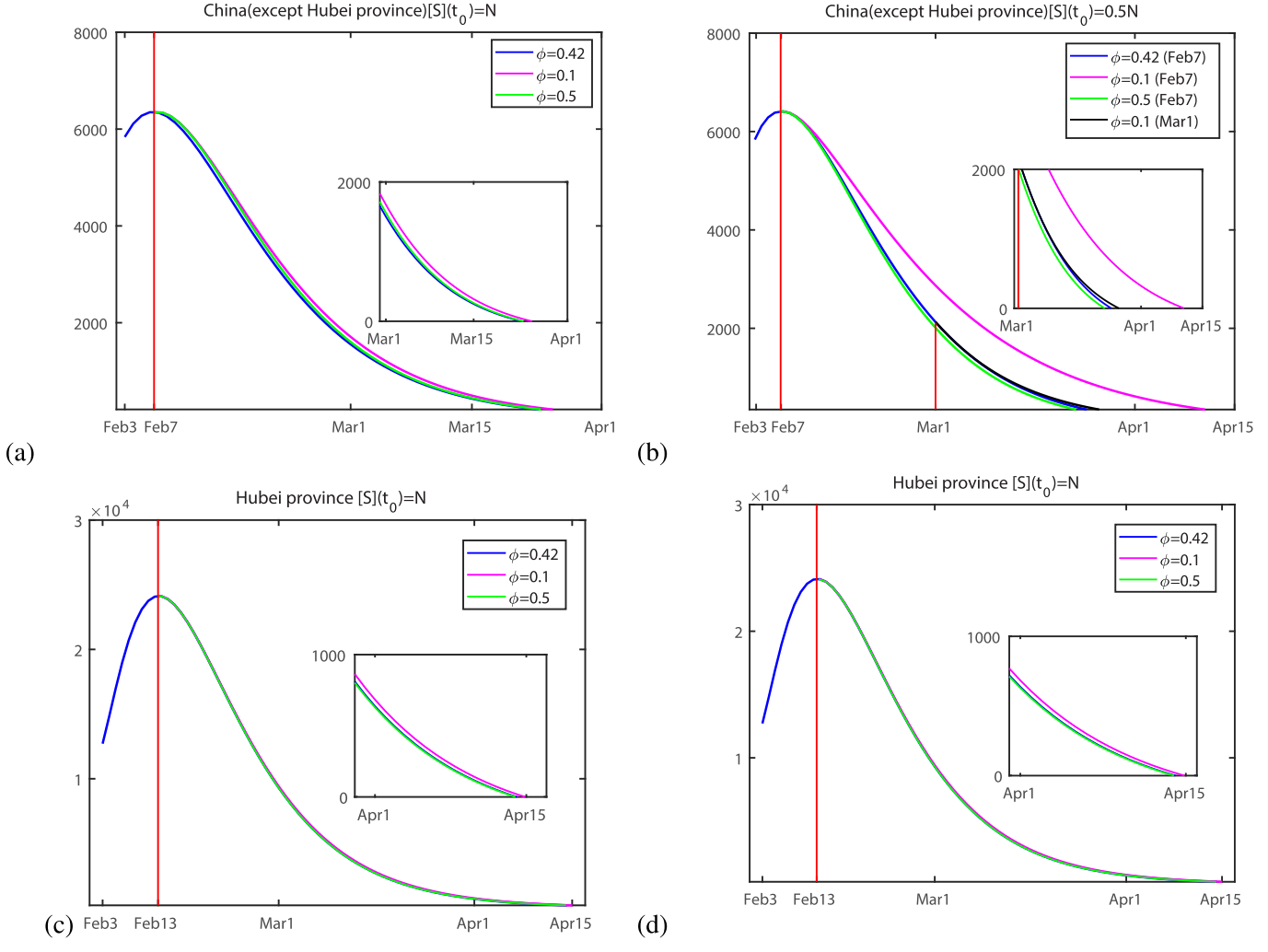

Fig. 9. Effects of the intensity of household quarantine measures (characterized by clustering coefficient φ) on COVID-19 spread in China except Hubei (top row: (a), (b)) and Hubei province (bottom row: (c), (d)) for different exposure proportions of susceptibles in the whole population: $[S](t_0)/N = 1$ (right column: (a), (c)), $[S](t_0)/N = 0.5$ (left column: (b), (d)). The red vertical lines in the plots mark different timings of adjusting quarantine measures.

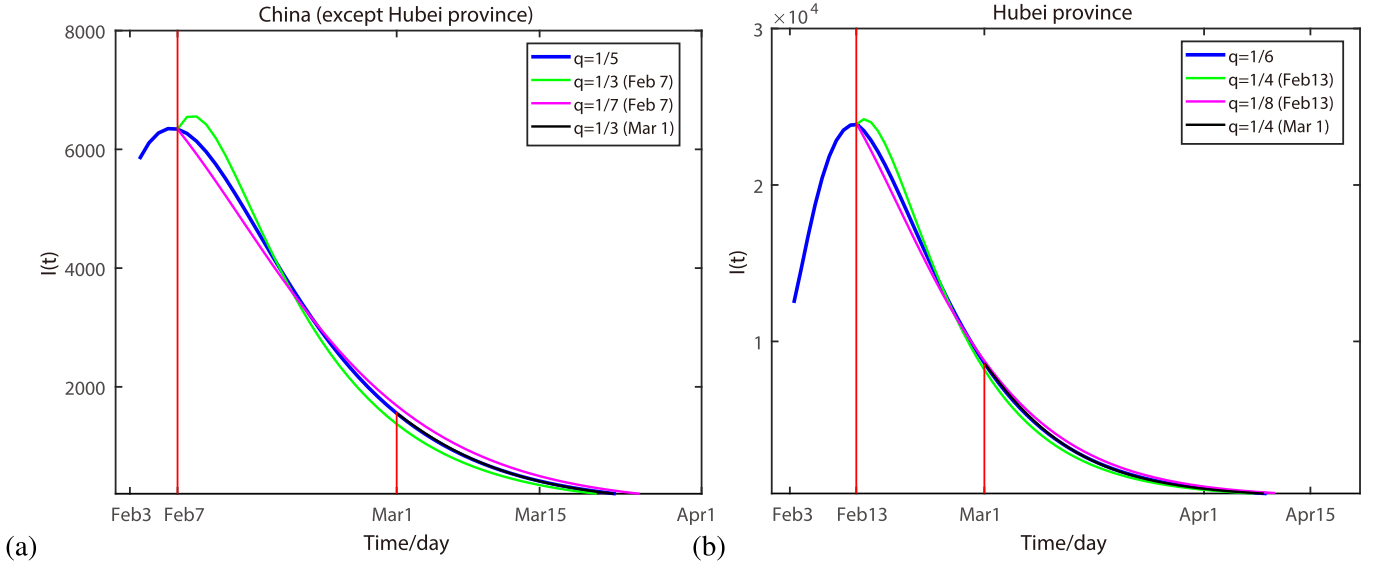

Fig. 10. The effects of different strength of detecting efforts on COVID-19 spread in (a) China except Hubei and (b) Hubei province. The red vertical lines in the plots mark different timings of adjusting quarantine measures.

Appendix A. Calculation of clustering coefficient

Since the closure of Wuhan city, Chinese villages and communities have basically been sealed off. Every family member basically stays at home every day. The ideal case is that every family is completely isolated. In this case, every family is a complete graph. According to the Chinese Statistic Yearbook (Chinese Statistic Yearboo, 2019), the number of family members varies from 1 person to 10 persons. The clustering coefficient φ of a network is defined as the ratio of the number of triangles to the number of triples in the network. This parameter can usually quantify the possibility that two friends of a person turn out to be friends of each other (Keeling et al., 1997; Rand, 1999). Therefore, in this special case of complete family quarantine, the clustering coefficient is maximal, which satisfies

$$\varphi = \frac{3\sum_{k=3}^{10} C_k^3 F_k}{\sum_{k=3}^{10} A_k^3 F_k} = \frac{\sum_{k=3}^{10}\left(\frac{k}{3}\right) F_k}{2\sum_{k=3}^{10}\left(\frac{k}{3}\right) F_k} = 0.5, \tag{A.1}$$

where F_k is the number of families each with k family members, C represents combination and A represents permutation.

When one family member goes out for living necessities, we let the number of links for every family increase by one. It is closest to the case of current household quarantine in China, and thus we consider this case in our paper. Ignoring the few addition of triangles, the clustering coefficient in this case is

$$\varphi = \frac{3\sum_{k=3}^{10} C_k^3 F_k}{\sum_{k=3}^{10} A_k^3 F_k + 2(k-1)} = \frac{\sum_{k=3}^{10}\left(\frac{k}{3}\right) F_k}{2\sum_{k=3}^{10}\left(\frac{k}{3}\right) F_k + 2(k-1)} = 0.42. \tag{A.2}$$

It follows from Eqs. (A.1) and (A.2) that the addition of inter-household connections (i.e., contacts between different families) leads to the decrease of the clustering coefficient in the entire population. The more inter-household connections, the smaller the clustering coefficient. In the special case when the household quarantine measure is revoked, the resumption of work and production will substantially increase individuals' outdoor connections, giving rise to a relatively smaller clustering coefficient than the value of $\varphi = 0.42$ in the case of household quarantine. Consider that every individual adds 7 outdoor contacts after they resume the work, then on average the total coefficient will be about $\varphi = 0.1$. In this paper, we adopted the case of $\varphi = 0.1$ to study the effects of resumption of work and production.

Appendix B. Elaborations of the pairwise epidemic model (1-14)

This subsection elaborates on the dynamical equations of pairwise epidemic model (1-14) presented in the Methods section 2.2. On the right hand side (rhs) of Eq. (1), the first term describes the addition of S individuals as the S_0 individuals are released from quarantine and return to the S state, the second and third terms denote the reduction of S individuals due to infection through contacts with E and A individuals, respectively. The fourth term corresponds to the reduction of S individuals as they are quarantined after being traced to have close contacts with confirmed cases of the disease. On the rhs of Eq. (2), the first and second terms represent the addition of E individuals due to infections caused by infected neighbors in states E and A, respectively. The third term depicts the outflow of E individuals after a lapse of latency period. The fourth term stands for the reduction of E individuals as they are quarantined by contact tracing. On the rhs of Eq. (3), the first and second terms stand for the inflow of I individuals from the E and E_0 states, respectively. The third term denotes the recovery of confirmed cases. On the rhs of Eq. (4), the first and second terms correspond to the inflow of A individuals from the E and E_0 individuals, respectively. The third term is the recovery of A individuals and the fourth term describes the reduction of A individuals due to quarantine by contact tracing. On the rhs of Eq. (5), the first, second and third terms represent the gains of R individuals due to recovery of I, A and A_0 individuals, respectively. On the rhs of Eq. (6), the first term depicts the quarantine of S individuals as a result of contact tracing and the second term describes the event that the S_0 individuals turn back to the S state after being released from quarantine. On the rhs of Eq. (7), the first term represents the quarantine of E individuals by contact tracing and the second term denotes the outflow of E_0 individuals after a quarantine period. On the rhs of Eq. (8), the first term stands for the inflow of A_0 resulting from quarantine of A individuals by contact tracing and the second term expresses the recovery of A_0 individuals.

On the rhs of Eq. (9), the first and second terms refer to the loss of SS-links due to infection through SSE-type triples and SSA-type triples, respectively. The third term considers the decrease of SS-links for the simple reason that one S individual is isolated from contact tracing. On the rhs of Eq. (10), the first and second terms describe the increase of SE-links as a result of infection through SSE-type triples and SSA-type triples, respectively. The third and fourth terms correspond to the reduction of SE-links as a result of direct infection and expiration of incubation period, respectively. The fifth and sixth terms depict the loss of SE-links because of infection through ESE-type and ASE-type triples, respectively. The seventh and eighth terms are attributed to quarantine of the S and E individuals, respectively, by contact tracing. On the rhs of Eq. (11), the first term denotes

the creation of *SA*-links because of the expiration of incubation period of the *E* individuals on *SE*-links. The second and third terms stand for the outflow of *SA*-links as a result of recovery and infection, respectively. The fourth and fifth terms refer to the loss of *SA*-links because of infection through *ESA*-type and *ASA*-type triples, respectively. The sixth and seventh terms represent the loss of *SA*-links as a result of quarantine of the *S* and *A* individuals, respectively, by contact tracing. On the rhs of Eq. (12), the first, second and third terms represent the addition of *EE*-links due to infection through the *SE*-links, *ESE*-type triples and *ASE*-type triples, respectively. The fourth and fifth terms account for the outflow of *EE*-links resulting from quarantine of one *E* individual on *EE*-links and *EEE*-type triples, respectively. On the rhs of Eq. (13), the first term considers the addition of *EA*-links because one of the *E* individuals on *EE*-links jump into the *A* class after the incubation period. The second, third and fourth terms describe the increase of *EA*-links because of infection through *SA*-links, *ESA*-type triples and *ASA*-type triples, respectively. The fifth and sixth term depict the loss of *EA*-links as a result of expiration of incubation period of *E* individuals and recovery of *A* individuals, respectively. The seventh and eighth terms refer to the outflow of *EA*-links due to quarantine of the *E* individuals and *A* individuals, respectively, by contact tracing. On the rhs of Eq. (14), the first term represents the creation of *AA*-links because the *E* individuals on the links jumps into the *A* class after the incubation period. The second term describes the loss of *AA*-links as one of the *A* individuals on the *AA*-link recovers. The third term considers the loss of *AA*-links due to quarantine of one of the *A* individuals by contact tracing.

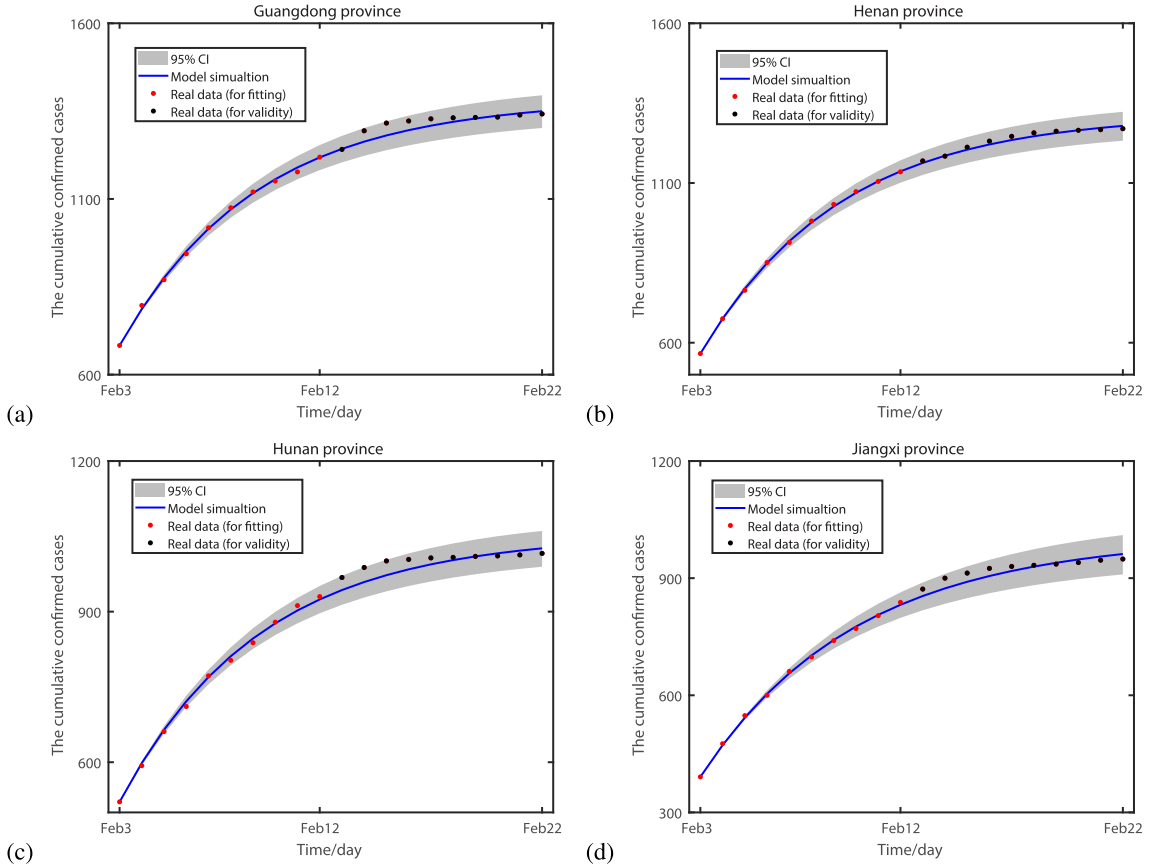

Fig. B.11. Fitting results of our model to real data of COVID-19 infections in other four hard-hit provinces: (a) Guangdong, (b) Henan, (c) Hunan, (d) Jiangxi. The red points are training data for parameter fitting, while the black points are real data for validation.

Appendix C. Fitting results in other four hard-hit provinces

To further calibrate the ability of our model in prediction of the epidemic behavior, we also present in Figure B11 the fitting results of COVID-19 spread in other four hard-hit provinces, namely Guangdong, Henan, Hunan and Jiangxi provinces. Indeed, the model has a good fit to the trajectory of the coronavirus prevalence in these regions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by the National Natural Science Foundation of China (grant numbers: 61873154, 12022113), and by the Shanxi Research Project on COVID-19 epidemic control and prevention (grant number: 202003D31011/GZ).

References

Chinese statistic Yearbook (in Chinese). http://www.stats.gov.cn/tjsj/ndsj/2019/indexch.htm, (2019).
Ali, S. T., Wang, L., Lau, E. H., Xu, X. K., Du, Z., Wu, Y., et al. (2020). Serial interval of SARS-CoV-2 was shortened over time by nonpharmaceutical interventions. *Science, 369*(6507), 1106–1109. https://doi.org/10.1126/science.abc9004
Anderson, S. C., Edwards, A. M., Yerlanov, M., et al. (2020). *Estimating the impact of COVID-19 control measures using a Bayesian model of physical distancing*. medRxiv. https://doi.org/10.1101/2020.04.17.20070086
Chan, K. H., & Yuen, K. Y. (2020). COVID-19 epidemic: Disentangling the re-emerging controversy about medical facemasks from an epidemiological perspective. *International Journal of Epidemiology, 49*(4), 1063–1066. https://doi.org/10.1093/ije/dyaa044
Eikenberry, S. E., Mancuso, M., Iboi, E., et al. (2020). To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infectious Disease Modelling, 5*, 293–308. https://doi.org/10.1016/j.idm.2020.04.001
Gamerman, D., & Lopes, H. F. (2006). *Markov chain Monte Carlo: Stochastic simulation for bayesian inference*. CRC Press.
Health Commission of Hubei Province. http://wjw.hubei.gov.cn/.
Hou, C., Chen, J., Zhou, Y., Hua, L., Yuan, J., He, S., et al. (2020). The effectiveness of quarantine of wuhan city against the corona virus disease 2019 (COVID-19): A well-mixed seir model analysis. *Journal of Medical Virology, 92*(7), 841–848. https://doi.org/10.1002/jmv.25827
House, T., & Keeling, M. J. (2011). Insights from unifying modern approximations to infections on networks. *Journal of The Royal Society Interface, 8*(54), 67–73. https://doi.org/10.1098/rsif.2010.0179
Hu, Z., Cui, Q., Han, J., Wang, X., Wei, E. I., & Teng, Z. (2020). Evaluation and prediction of the COVID-19 variations at different input population and quarantine strategies, a case study in Guangdong province, China. *International Journal of Infectious Diseases, 95*, 231–240. https://doi.org/10.1016/j.ijid.2020.04.010
Keeling, M. J. (1999). The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 266*(1421), 859–867. https://doi.org/10.1098/rspb.1999.0716
Keeling, M. J., Rand, D. A., & Morris, A. J. (1997). Correlation models for childhood epidemics. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 264*(1385), 1149–1156. https://doi.org/10.1098/rspb.1997.0159
Leung, K., Wu, J. T., Liu, D., & Leung, G. M. (2020). First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: A modelling impact assessment. *The Lancet, 395*(10233), 1382–1393. https://doi.org/10.1016/S0140-6736(20)30746-7
Li, M. Y. (2018). *An introduction to mathematical modeling of infectious diseases* (Vol. 2). Cham, Switzerland: Springer.
Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., et al. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus infected pneumonia. *New England Journal of Medicine*. https://doi.org/10.1056/NEJMoa2001316
Luo, X., & Jin, Z. (2020). A new insight into isolating the high-degree nodes in network to control infectious diseases. *Communications in Nonlinear Science and Numerical Simulation, 91*, 105363. https://doi.org/10.1016/j.cnsns.2020.105363
Newman, M. E. (2010). *Newtworks: An introduction*. New York.
Official news on the change of diagnosed approaches issued by the Chinese govenment (in Chinese) http://www.gov.cn/zhengce/zhengceku/2020-02/09/content_5476407.htm.
Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical Review Letters, 86*(14), 3200. https://doi.org/10.1103/PhysRevLett.86.3200
The platform of 2019-nCov infection. http://www.clas.ac.cn/xwzx2016/163486/xxfysjpt2020/.
Rand, D. A. (1999). Correlation equations and pair approximations for spatial ecologies. *Advanced Ecological Theory: Principles and Applications*, 100–142. https://doi.org/10.1002/9781444311501
Roda, W. C., Varughese, M. B., Han, D., & Li, M. Y. (2020). Why is it difficult to accurately predict the COVID-19 epidemic? *Infectious Disease Modelling, 5*, 271–281. https://doi.org/10.1016/j.idm.2020.03.001
Shao, P., & Shan, Y. G. (2020). *Beware of asymptomatic transmission: Study on 2019-nCov preventtion and control measures based on SEIR model*. https://doi.org/10.1101/2020.01.28.923169
Shen, M. W., Peng, Z. H., Xiao, Y. N., & Zhang, L. (2020). *Modelling the epidemic trend of the 2019 novel coronavirus outbreak in China*. https://doi.org/10.1101/2020.01.23.916726
Tang, B., Bragazzi, N. L., Li, Q., Tang, S., Xiao, Y., & Wu, J. (2020a). An updated estimation of the risk of transmission of the novel coronavirus (2019-nCov). *Infectious Disease Modelling, 5*, 248–255. https://doi.org/10.1016/j.idm.2020.02.001
Tang, B., Wang, X., Li, Q., Bragazzi, N. L., Tang, S., Xiao, Y., et al. (2020b). Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. *Journal of Clinical Medicine, 9*(2), 462. https://doi.org/10.3390/jcm9020462
Tang, B., Xia, F., Tang, S., Bragazzi, N. L., Li, Q., Sun, X., et al. (2020c). The effectiveness of quarantine and isolation determine the trend of the COVID-19 epidemics in the final phase of the current outbreak in China. *International Journal of Infectious Diseases, 95*, 288–293. https://doi.org/10.1016/j.ijid.2020.03.018
The National Health Commission of P. R. China. http://www.nhc.gov.cn/.
Van den Driessche, P., & Watmough, J. (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences, 180*(1–2), 29–48. https://doi.org/10.1016/S0025-5564(02)00108-6
Wang, C. J., Ng, C. Y., & Brook, R. H. (2020). Response to COVID-19 in taiwan: Big data analytics, new technology, and proactive testing. *Journal of the American Medical Association, 323*(14), 1341e1342. https://doi:10.1001/jama.2020.3151.
World Health Organization (WHO). https://covid19.who.int/.
Wu, P., Hao, X., Lau, E. H., Wong, J. Y., Leung, K. S., Wu, J. T., et al. (2020). Real-time tentative assessment of the epidemiological characteristics of novel coronavirus infections in Wuhan, China, as at 22 January 2020. *Euro Surveillance, 25*(3), 2000044. https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000044
Xue, L., Jing, S., Miller, J. C., Sun, W., Li, H., Estrada-Franco, J. G., et al. (2020). A data-driven network model for the emerging COVID-19 epidemics in Wuhan, Toronto and Italy. *Mathematical Biosciences, 326*, 108391. https://doi.org/10.1016/j.mbs.2020.108391
Yang, J., & Xu, F. (2019). The computational approach for the basic reproduction number of epidemic models on complex networks. *IEEE Access, 7*, 26474–26479. https://doi.org/10.1109/ACCESS.2019.2898639
Yang, Z., Zeng, Z., Wang, K., Wong, S. S., Liang, W., Zanin, M., et al. (2020). Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *Journal of Thoracic Disease, 12*(3), 165. https://doi.org/10.21037/jtd.2020.02.64
You, C., Deng, Y., Hu, W., Sun, J., Lin, Q., Zhou, F., et al. (2020). Estimation of the time-varying reproduction number of COVID-19 outbreak in China. *International Journal of Hygiene and Environmental Health, 228*, 113555. https://doi.org/10.1101/2020.02.08.20021253