


Article

Analysis of potential risk of COVID-19 infections in China based on a pairwise epidemic model

Xiaofeng Luo ^{1,2} , Shanshan Feng ^{1,3}, Junyuan Yang ^{1,2}, Xiao-Long Peng ^{1,2}, Xiaochun Cao ^{1,2}, Juping Zhang ^{1,2}, Meiping Yao ⁴, Huaiping Zhu ⁵, Michael Y. Li ⁶, Hao Wang ⁶ and Zhen Jin ^{1,2,*}

¹ Complex Systems Research Center, Shanxi University, Taiyuan 030006, Shanxi, P.R. China

² Shanxi Key Laboratory of Mathematical Techniques and Big Data Analysis on Disease Control and Prevention, Shanxi University, Taiyuan 030006, Shanxi, P.R. China

³ School of Data Science and Technology, North University of China, Taiyuan 030051, Shanxi, P.R. China

⁴ School of Mathematical Sciences, Shanxi University, Taiyuan 030006, Shanxi, P.R. China

⁵ Department of Mathematics and Statistics, York University, Toronto, ON M3J 1P3, Canada

⁶ Department of Mathematics and Statistics Sciences, University of Alberta, Edmonton, Alberta, T6G 2G1, Canada

* Correspondence: jinzhn@263.net

Abstract: The ongoing outbreak of the novel coronavirus pneumonia (also known as COVID-19) has triggered a series of stringent control measures in China, such as city closure, traffic restrictions, contact tracing and household quarantine. These containment efforts often lead to changes in the contact pattern among individuals of the population. Many existing compartmental epidemic models fail to account for the effects of contact structure. In this paper, we devised a pairwise epidemic model to analyze the COVID-19 outbreak in China based on confirmed cases reported during the period February 3rd–17th, 2020. By explicitly incorporating the effects of family clusters and contact tracing followed by household quarantine and isolation, our model provides a good fit to the trajectory of COVID-19 infections and is useful to predict the epidemic trend. We obtained the average of the reproduction number $R = 1.494$ (95% CI: 1.483 – 1.507) for Hubei province and $R = 1.178$ (95% CI: 1.145 – 1.158) for China (except Hubei), suggesting that some existing studies may have overestimated the reproduction number by neglecting the dynamical correlations and clustering effects. We forecasted that the COVID-19 epidemic would peak on February 13th (95% CI: February 9 – 17th) in Hubei and 6 days earlier in the regions outside Hubei. Moreover the epidemic was expected to last until the middle of March in China (except Hubei) and late April in Hubei. The sensitivity analysis shows that ongoing exposure for the susceptible and population clustering play an important role in the disease propagation. With the enforcement of household quarantine measures, the reproduction number R effectively reduces and epidemic quantities decrease accordingly. Furthermore, we gave an answer to the public concern on how long the stringent containment strategies should maintain. Through numerical analysis, we suggested that the time for the resumption of work and production in China (except Hubei) and Hubei would be the middle of March and the end of April, 2020, respectively. These constructive suggestions may bring some immeasurable social-economic benefits in the long run.

Keywords: COVID-19; pairwise epidemic model; household quarantine; clustering coefficient

1. Introduction

A zoonotic coronavirus (named as COVID-19), firstly detected in Wuhan, a central and crowded city of China, has crossed species to infect human population [1,2]. The symptoms of COVID-19 include cough, fever and difficulty breath, most like SARS (Severe acute respiratory syndrome coronavirus) and MERS (Middle East respiratory syndrome coronavirus) [3]. The period for such symptoms from mild to severe respiratory infections lasts 2 – 14 days. As indicated, the transmission routes contain direct transmission, such as touching or shaking hands and indirect transmission consist of the

air by coughing and sneezing, even if contacting some contaminated things by virus particles [3]. Additionally, coronaviruses can be extremely contagious and spread easily from person to person. To date, cases of the COVID-19 coronavirus have been confirmed more than 70,000 - mostly in China, as well as in Japan, South Korea, Thailand, Singapore, Philippines, and the United States of America etc, and with nearly 2000 reported deaths. Apparently, outbreaks of COVID-19 have become a globally public health concern in medical community as the virus is spreading around the world.

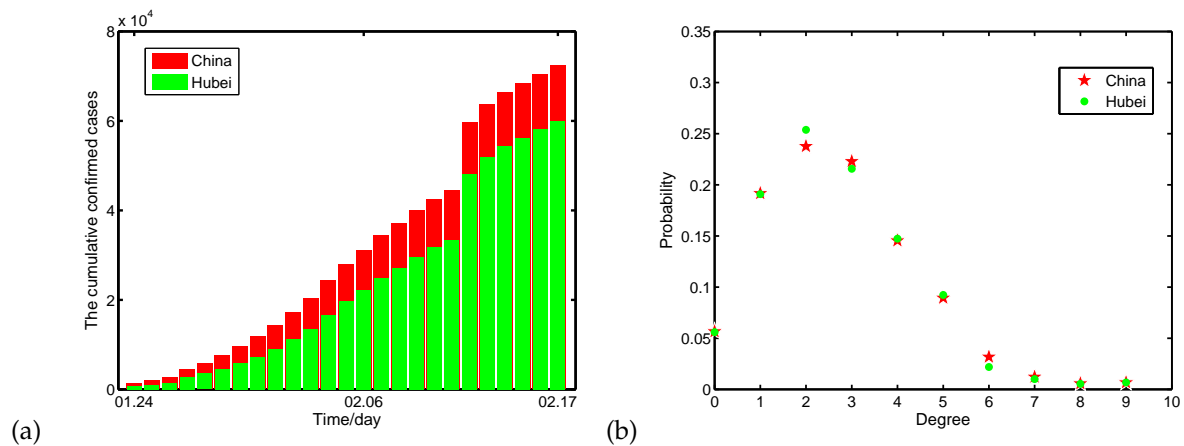


Figure 1. (a) Real data of cumulative confirmed cases in China (red) and Hubei province (green). (b) Degree distributions for China and Hubei province.

The surge of confirmed cases continues even though some control measures such as city closure, extended holidays, travel restrictions and closure of live trade markets have been taken. Some reasons for the rapid growth of cases include increased travel for Chinese New Year, the long incubation period, increased detection, asymptomatic transmission, substantial person-to-person transmission, and continued environmental or animal exposure to a source of infection. As of 24pm on February 17th, 2020, the number of confirmed cumulative cases rose to 72436. The top five provinces in China are Hubei, Guangdong, Henan, Jiangxi, and Hunan, respectively. The reported cases of COVID-19 infections in Hubei province account for about 95% of the infections across China (see Figure1(a)). In order to cut off COVID-19 infection source, Chinese government took an extreme control strategy-prohibiting all the transports in and out of Wuhan and advised all the Chinese citizens to isolate themselves at home at 10am on January 23rd, 2020 [4]. However, more than 5 million people have left Wuhan during the Spring Festival period before January 23rd, 2020, which increased the risk of COVID-19 infections [5]. The closure of Wuhan changed the behaviors of all the Chinese. As a result, before and after the city closure, there is a substantial change on the pattern of the COVID-19 infections. How to evaluate such a switch strategy play an important role in government decision- and policy- makers. Mathematical modelling is a useful tool to describe mechanisms of epidemic transmission and estimate the efficacy of some control strategies.

Although there are many mathematical models to study the dynamics of COVID-19 infection in China [6–12], they always assume that individuals are homogeneously mixed and ignore the heterogeneous contact structure between individuals. In fact, under such an extreme control strategy every person should almost stay at home and substantially reduce their contacts with outside. According to Chinese Statistic Yearbook [13], there are ten composition forms in the household structure. In network terminology, the population can be regarded as a contact network consisting of nodes and links, in which the nodes mimic individuals and links stand for social contacts between individuals among the population. The contact pattern is often characterized by a node degree distribution, where the node degree is the number of contacts an individual has. The contact structure under household quarantine (in this scenario every individual only contacts with its family members) exhibits a unimodal degree distribution (see Figure1(b)). In addition, such contact patterns under

household quarantine exhibits strong clustering effects, depicting the closeness of triple cliques in a network. However, the widely used compartmental modelling approaches typically fail to characterize the effects of clustering feature, heterogeneity in contact behavior and dynamical correlations on COVID-19 propagation. In this paper, we devised a pairwise epidemic model to investigate the COVID-19 epidemics based on confirmed cases reported during the period February 3rd–17th, 2020. To our knowledge, many researchers have paid much attention on closure schemes for pairwise approximation for intensive computations and theoretical analysis [14–18]. Since such generated network belongs to a clustering with a homogeneous degree, we took the moment closure approach proposed by Keeling *et al.* [19,20] to approximate the number of triples.

For each emerging infectious disease, some common concerns include the reproduction number, transmission time interval, peak time, peak values and final size. These quantities determine the potential and severity of an outbreak and are in favor of evaluation of control measures and providing a guidance for slowing down the disease propagation. In this project, we based on data of COVID-19 infections collected from National Health Commission of China (NHCC) [21] to identify the parameters of the proposed model and calculated all the mentioned quantities. In particular, we forecasted the transmission dynamics of COVID-19 in Hubei province as well as in regions outside Hubei. 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. 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. In this work we investigated the effects of the city closure and close contact tracing followed by household quarantine. To our best knowledge, there are few works coupling the data and epidemic models on complex networks to research a specific disease. We tried to build a bridge connecting the two aspects to realize the theorem and practice.

2. Materials and Methods

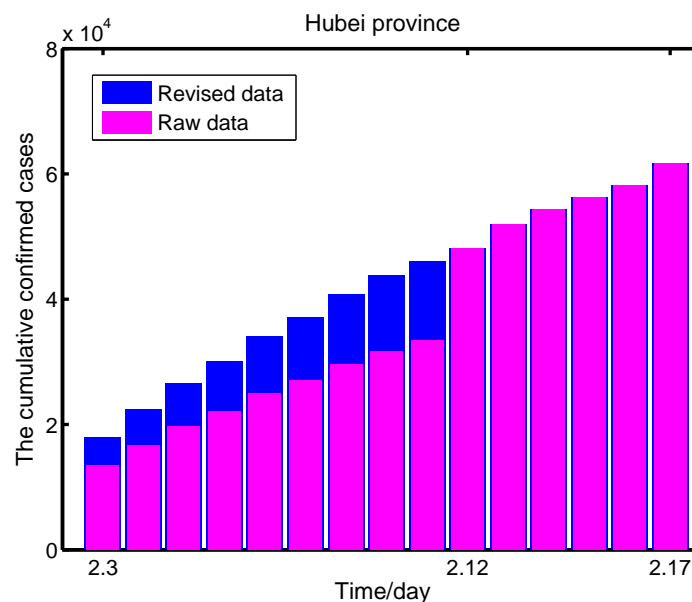


Figure 2. Comparison between reported data and revised data of cumulative confirmed cases of COVID-19 in Hubei province.

2.1. Data collection and description

We collected all the daily reported data on COVID-19 infections 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 [21,22].

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 [23]—there exists a singular point for cumulative confirmed cases on February 12th, 2020 (see Figure 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 Figure 2 for the revised data).

The clustering coefficient denoted by ϕ , as one of the network characteristic parameters, can 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 [13], we obtained that when every family is completely isolated, $\phi = 0.5$ and when only one family member is allowed to go out for living necessities, $\phi = 0.42$ (see Appendix A for details).

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 network 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.

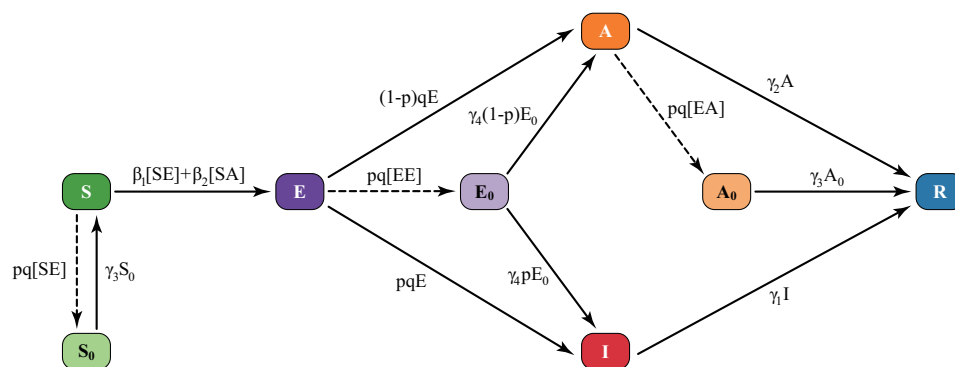


Figure 3. The flow diagram of pairwise model for COVID-19 epidemic.

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 Figure 3, we classified all of the individuals (nodes) into eight different states: unquarantined susceptible (S), quarantined susceptible (S_0), unquarantined latent (E), quarantined latent (E_0), unquarantined asymptomatic (A), quarantined asymptomatic (A_0), confirmed infected (I) and recovered (R). 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 latent 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 recovery rate γ_1 and γ_2 , 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 $1/\gamma_4$ 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 Figure 3. Parameters used in the model have been summarized in Table 2.

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 exposed (latent) individuals |
| $[E_0]$ | The number of quarantined exposed (latent) individuals (i.e., the number of latent 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 |

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.

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

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

$$\frac{d[I](t)}{dt} = pq[E] + p\gamma_4[E_0] - \gamma_1[I], \quad (3)$$

$$\frac{d[A](t)}{dt} = (1-p)q[E] + (1-p)\gamma_4[E_0] - \gamma_2[A] - pq[EA], \quad (4)$$

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

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

$$\frac{d[E_0](t)}{dt} = pq[EE] - \gamma_4[E_0], \quad (7)$$

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

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.

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

$$\begin{aligned} \frac{d[SE]}{dt} &= \beta_1[SSE] + \beta_2[SSA] - \beta_1[SE] - q[SE] \\ &\quad - \beta_1[ESE] - \beta_2[ASE] - pq[SEE] - pq[ESE], \end{aligned} \quad (10)$$

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

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

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

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

For detailed explanation of the biological meaning of Eqs. (1-14), the readers are referred to Appendix 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.* [19,20] as follows:

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

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 [25] or the next generation matrix method [26], the expression R takes the form of

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

where

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

and

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

Since the pairwise epidemic model (1-14) is a networks 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 $(1 - \phi)(n - 1)$ 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 [28], 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

$$\bar{z}(T) = \sum_{t=1}^T \left(pq[E](t) + p\gamma_4[E_0](t) \right), \quad T = 1, 2, \dots, 15, \quad (19)$$

and its variation with a small perturbation by

$$z(T) = \bar{z}(T) + \delta, \quad T = 1, 2, \dots, 15, \quad (20)$$

where $z(1) = \bar{z}(1)$, $T = 1, 2, \dots, 15$ corresponds to the days during the period February 3rd–17th, 2020, and δ is supposed to follow the normal distribution $N(0, \varepsilon^2)$ with $\varepsilon = 1000$. Given the parameter vector Θ , the maximum likelihood function for the revised data $Z = \{z(T)\}_{T=0}^{15}$ is

$$L(Z|\Theta) = \prod_{T=1}^{15} \frac{1}{\sqrt{2\pi\varepsilon}} \exp \left(-\frac{[z(T) - \bar{z}(T)]^2}{\varepsilon^2} \right). \quad (21)$$

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 [28], 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

$$\min \left\{ 1, \frac{P(\tilde{\Theta}|Z)}{P(\Theta|Z)} \right\}. \quad (22)$$

2.5. Parameters and initial conditions

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

| Parameter | Definition | China (except Hubei) | Standard deviation | Hubei | Standard deviation | Data source |
|--------------|---|----------------------------|-----------------------|------------|-----------------------|----------------|
| β_1 | transmission rate by E individuals | 0.158 | 0.046 | 0.271 | 0.079 | MCMC |
| β_2 | transmission rate by A individuals | 0.114 | 0.031 | 0.267 | 0.061 | MCMC |
| p | probability of showing symptoms | 0.751 | 0.039 | 0.704 | 0.041 | MCMC |
| $1/q$ | latent period for E individuals | 6 | – | 7 | – | [6] |
| γ_1 | The recovery rate of I individuals | 1/10 | – | 1/10 | – | [21,22] |
| γ_2 | The recovery rate of A individuals | 1/12 | – | 1/12 | – | [21,22] |
| $1/\gamma_3$ | quarantine period for S and A individuals | 14 | – | 14 | – | [6] |
| $1/\gamma_4$ | quarantine period for E individuals | 14 | – | 14 | – | [6] |
| N | total population | 1,336,210,000 | – | 59,170,000 | – | [13] |
| ϕ | clustering coefficient | 0.42 | – | 0.42 | – | [13] |
| n | average node degree | 3.1 | – | 3.04 | – | [13] |

The population size of China (except Hubei) and Hubei province is, respectively, about 1336210000 and 59170000 [13], so we set $[S](t_0)$ to be 1336210000 and 59170000 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 13522 and 6028, 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 12712 ($= 13522 - 810$) and 5837 ($= 6028 - 169$) 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 $\gamma_3 = \gamma_4 = 1/14$. According to literature [6], γ_2 is about 1/12, and the latent period is about 7 days. However, as the test technique and supply of medical resources improve, especially outside Hubei province, we set the latent period $1/q$ to 6 days for China (except Hubei) and 7 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 (Figure 1(b)) obtained from the data published on the Chinese Statistic Yearbook [13], 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,

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

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 the formula (23) gives a real number less than 1.

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

Table 3. Initial values used in the model.

| Initial values | Definition | China (except Hubei) | Standard deviation | Hubei | Standard deviation | Data source |
|----------------|---------------------------------------|----------------------|--------------------|------------|--------------------|-------------|
| $[S](t_0)$ | initial number (#) of S individuals | N | – | N | – | [13] |
| $[E](t_0)$ | initial # of E individuals | 6,550 | 90.4 | 28,826 | 225.9 | MCMC |
| $[I](t_0)$ | initial # of I individuals | 5,837 | – | 12,712 | – | [21,22] |
| $[A](t_0)$ | initial # of A individuals | 1,091 | 23.9 | 10,705 | 124.9 | 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 | 1,925 | 36.8 | 21,577 | 213.9 | MCMC |
| $[SA](t_0)$ | initial # of SA -links | 1,971 | 20.1 | 23,610 | 374.7 | MCMC |
| $[SS](t_0)$ | initial # of SS -links | 1,494,823,098 | 21,091,366 | 47,903,399 | 351,255 | 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) |

3. Results

3.1. Fitting results

Based on the analysis of the diagnosed cases of COVID-19 infection reported in [4], we adopted the MCMC method [28] 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. Figure 4 shows a good fitting between the model solution and real data, well suggesting the epidemic trend in China under the intervention of the home-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 $R = 1.494$ with 95% CI of [1.483, 1.507] for the Hubei province and $R = 1.178$ with 95% CI of [1.145, 1.258] for the regions outside Hubei province. 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 Figure A1 in Appendix C). In what follows, we mainly employed our model with estimated parameters to explore the epidemic behavior of COVID-19 infections in China (except Hubei) and in Hubei.

3.2. Prediction of epidemiological quantities

Applying the estimated parameter values, we forecasted that the outbreak of COVID-19 in China (except Hubei province) is expected to peak on February 7th, 2020 (95% CI between February 3rd and February 11st) and the peak value in terms of confirmed infectious cases is predicted to be $I_{max} = 6,253$ with 95% CI of [6, 150, 6, 368] (Figure 5(a)), whereas the peak time of COVID-19 infections in Hubei province is expected to be on February 13th, 2020 (95% CI between February 9th and February 17th) with the peak value of $I_{max} = 24,674$ (95% CI [24, 314, 24, 582]) (Figure 5(b)). In addition, the COVID-19 epidemic in China (except Hubei) and Hubei province is expected to terminate by the end of March and late April, respectively. The final sizes are forecasted to be 13,757, with 95% CI of [12, 678, 14, 752] in China (except Hubei) and 74,057 with 95% CI of [73, 351, 74, 755] in Hubei province. The primary reason for the COVID-19 epidemic ending late with a larger final size in Hubei province is that the

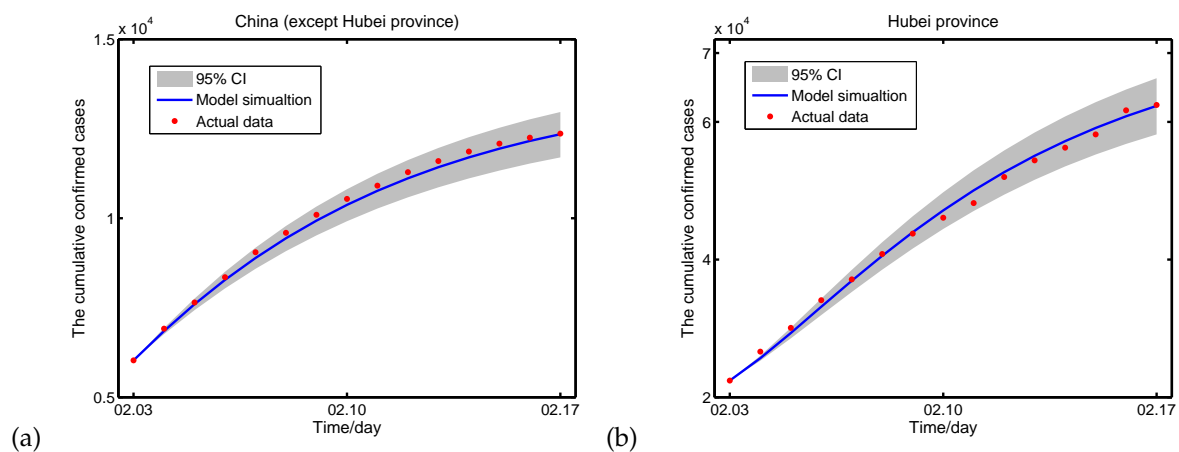


Figure 4. The fitting results of our model to real data of COVID-19 infections in China exclusive of Hubei province (a) and that in Hubei province (b). The grey area marks the 95% CI of MCMC estimations.

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 is worth noting that the epidemiological predictions presented in this paper are based on the current control strategies. Arguably, the peak time, peak value and duration of the epidemic will change accordingly as long as the disease control measures are modified.

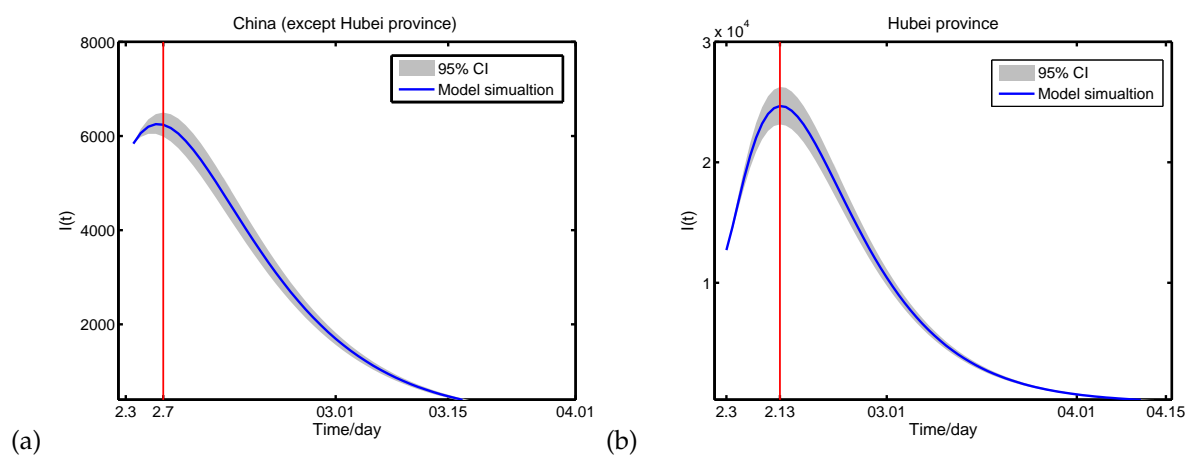


Figure 5. The time series of the number $I(t)$ of confirmed individuals individuals in China except Hubei province (a) and Hubei province (b). The red vertical lines in the plots mark the peak time for the COVID-19 spread of which the trajectory is plotted in blue curves.

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.

Figure 6 and Tables 4-5 show that shortening the exposed period $1/q$ from 8 days to 4 days advances the peak arrival time and decreases the final size, although increasing the peak size of confirmed cases. Also, the epidemic duration advances about one week. Increasing the detection rates as an increase of p is in favor of mitigating COVID-19 transmission.

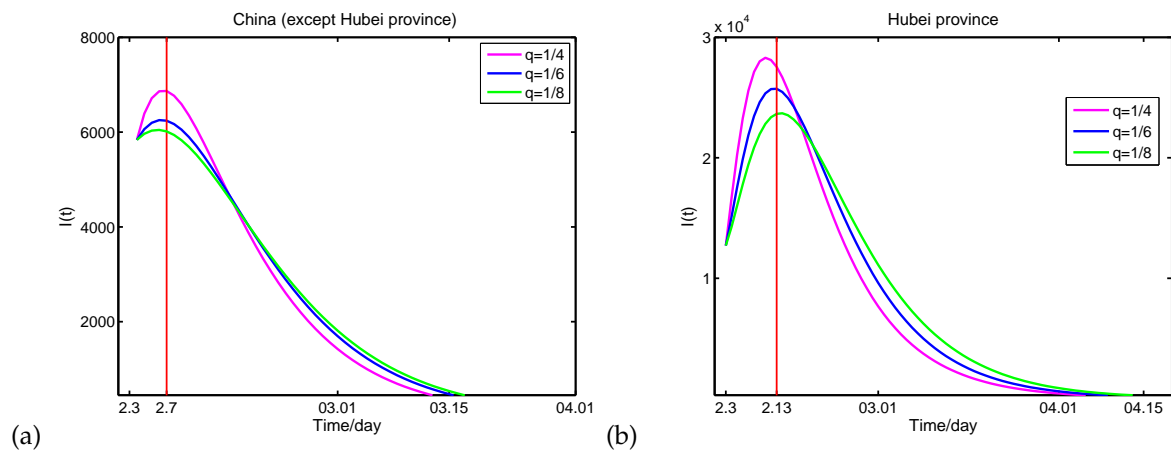


Figure 6. The numbers of confirmed infected individuals change with different latent periods $1/q$ for: (a) China (except Hubei); (b) Hubei province. The red vertical lines in the plots mark the peak time for the COVID-19 spread of which the trajectory is plotted in blue curves.

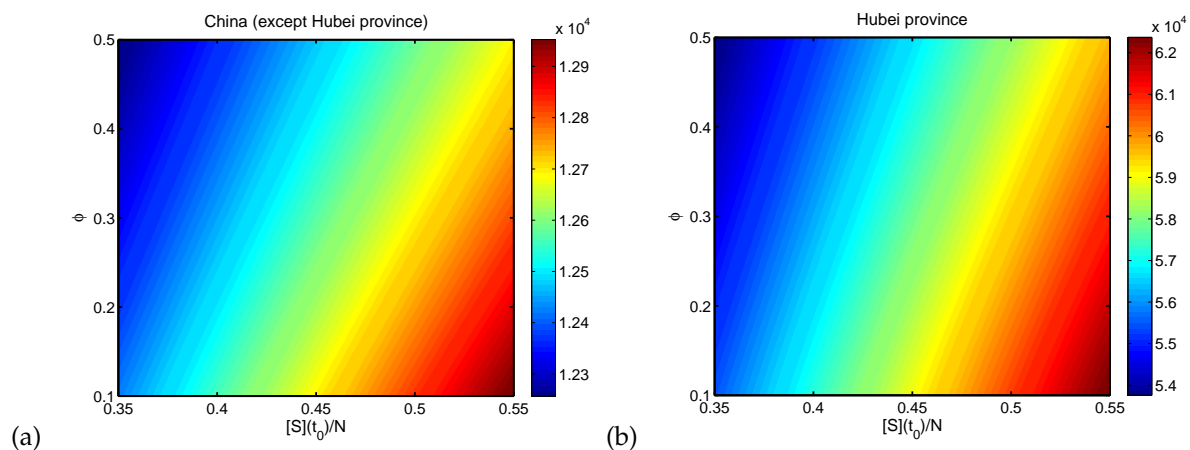


Figure 7. Contour plots with regard to the clustering coefficient ϕ and the initial exposure proportion of susceptibles $[S](t_0)/N$ in the whole population for (a) China except Hubei province and (b) Hubei province.

As Section 2.1 introduced, the intensity of isolation under the household quarantine measure is characterized by the clustering coefficient ϕ . Table 4-5 shows that, with the increase of ϕ or the intensity of isolation in China, the peak size and the final 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. Figure 7 presents the contour figure of the final 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 final sizes increase accordingly. In next section, we will further discuss their effects on COVID-19 combined with control measures.

4. Discussion

COVID-19 is rife in China. The Chinese government has taken many extreme strategies after January 23rd, 2020, which resulting in strong correlation between exposed individuals, asymptomatic individuals, confirmed cases and susceptible people. Moreover, 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 will be quarantined somewhere (at

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

| Parameters | Peak size | Peak time | Final size | Epidemic duration |
|--------------|-----------|------------|------------|-------------------|
| $q=1/4$ | 6,868 | 2020.02.06 | 13,477 | 42 days |
| $q=1/6$ | 6,253 | 2020.02.07 | 13,757 | 45 days |
| $q=1/8$ | 5,902 | 2020.02.08 | 13,975 | 48 days |
| $\phi = 0.1$ | 6,256 | 2020.02.07 | 14,154 | 47 days |
| $\phi = 0.3$ | 6,254 | 2020.02.07 | 13,893 | 46 days |
| $\phi = 0.5$ | 6,252 | 2020.02.07 | 13,673 | 45 days |

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

| Parameters | Peak size | Peak time | Final size | Epidemic duration |
|--------------|-----------|------------|------------|-------------------|
| $q=1/4$ | 28,302 | 2020.02.11 | 70,468 | 66 days |
| $q=1/6$ | 25,725 | 2020.02.13 | 73,070 | 70 days |
| $q=1/8$ | 23,694 | 2020.02.14 | 74,876 | 74 days |
| $\phi = 0.1$ | 25,193 | 2020.02.14 | 77,742 | 75 days |
| $\phi = 0.3$ | 24,911 | 2020.02.13 | 75,439 | 73 days |
| $\phi = 0.5$ | 24,509 | 2020.02.13 | 73,182 | 72 days |

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.

At the time of writing, what people concern is when they can return to work and school. 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, $\phi = 0.42$ represents the gathering degree in current household quarantine measure. Once returning to work and school, people interconnects closely and the clustering coefficient decreases (see Appendix A). Meanwhile, Figure 7 indicates isolation measure may change ϕ and the exposure proportion of susceptible individuals. Therefore, Figure 8 provides some constructive advise based on our network model. We have seen from Figure 8(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 7th February, the final 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 final size increases from 13,757 to 17,791 and the epidemic duration delays from 79 days to 100 days, and when the same change of ϕ appears on March 1st, the final size and epidemic duration changes slightly (see Figure 8 (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 the middle of March. Similarly, in Hubei province, when increasing ϕ from 0.42 to 0.5 or decreasing ϕ from 0.42 to 0.1 on the 13th February in terms of total population (see Figure 8 (c)), the final size and the epidemic duration have trivial changes, whereas when taking the same changes in terms of partial population (Figure 8 (d)), the final size increases from 74,055 to 98,593 and the epidemic duration delays from 106 to 119 days. Thus, it may be plausible to resume work, production and other gathering activities in Hubei province until the end of April.

The q , one of our model parameters, reflects the strength of testing efforts. Figure 9(a) indicates that when intensifying the medical detection (q from $1/8$ to $1/4$) on February 7th, 2020 in China (except Hubei), the final size decreased from 13,859 to 13,625 and the epidemic duration advances about one week. However, when taking the same detection change on 1st March, the effects are small. Likewise, similar impacts happen in Hubei province. It means that the medical detection is promoted earlier, COVID-19 ends earlier and the final size is smaller.

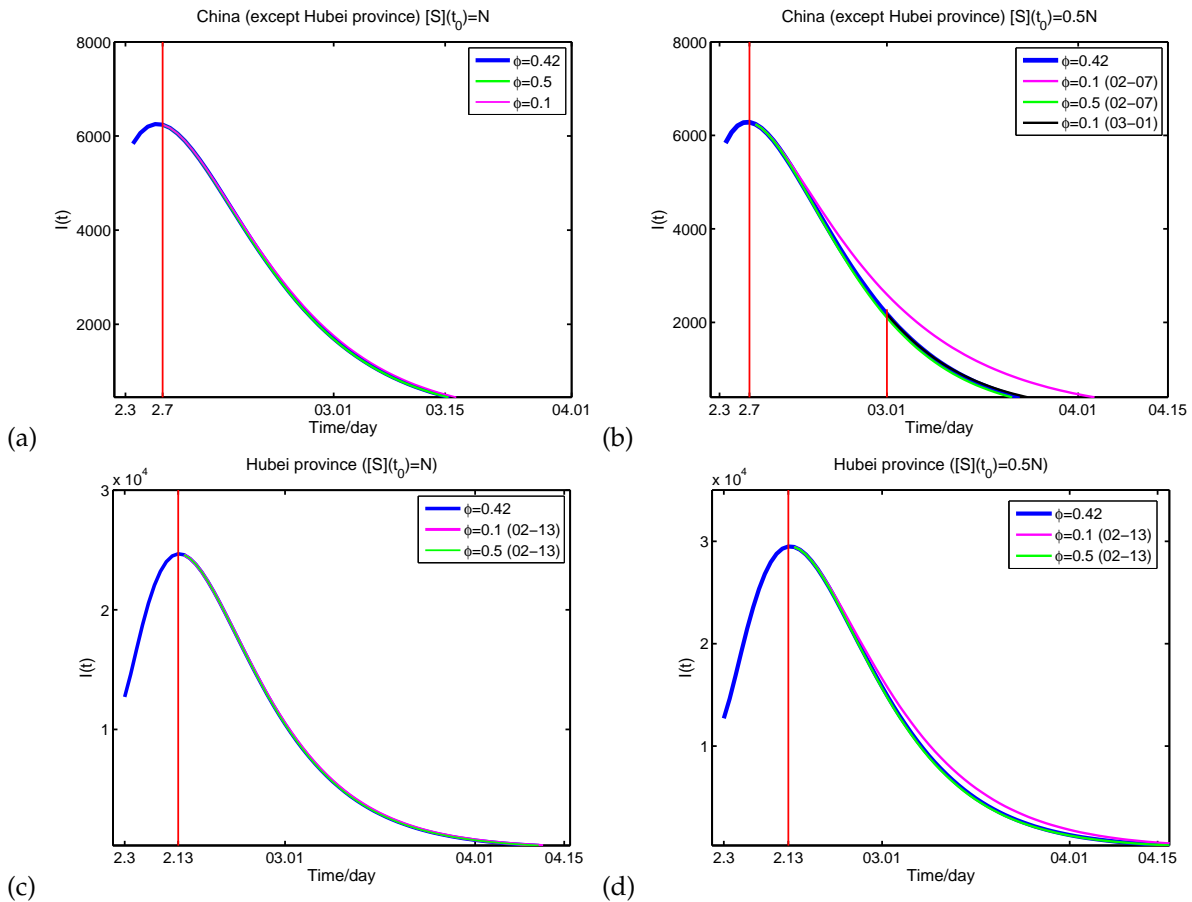


Figure 8. Effects of the intensity of household quarantine measures (characterized by clustering coefficient ϕ) on COVID-19 spread in China except Hubei (top row: (a), (b)) and Hubei province (bottom row: (c), (d)) for different exposure proportions of susceptibles in the whole population: $[S](t_0)/N = 1$ (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

5. Conclusion

An emerging outbreak of COVID-19 in China has lasted almost three months and it brings about 77,046 confirmed cases, 2,445 deaths up to February 23rd, 2020. 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 exhibits in form of family clusters and hence the transmission patterns of COVID-19 infections have radically altered. Considering such signatures, we proposed a pairwise epidemic model on a network with relevant population structure to address the dynamics of COVID-19 propagation in China. Coupling surveillance data started on February 3rd with our proposed model, the reproduction numbers are evaluated as 1.494 (95% CI [1.483, 1.507]) for Hubei province and 1.178 (95% CI [1.145, 1.258]) for China except Hubei, which are obviously smaller than those published in some literature [6,8,11,12] and they seem to be consistent with the result in [2]. Actually, there are two primary reasons: one is that those values of estimated parameters in the model heavily depend on the choice of the initial surveillance data, the other is that most of existing results ignore clustering effects induced by home isolation as well as dynamical correlations between susceptible individuals and infectious individuals. Interestingly, our model successfully mimics the time surveillance data from February 3rd to February 17th and the solution of model in China (see Figure 4). Moreover, our results can be extended to make

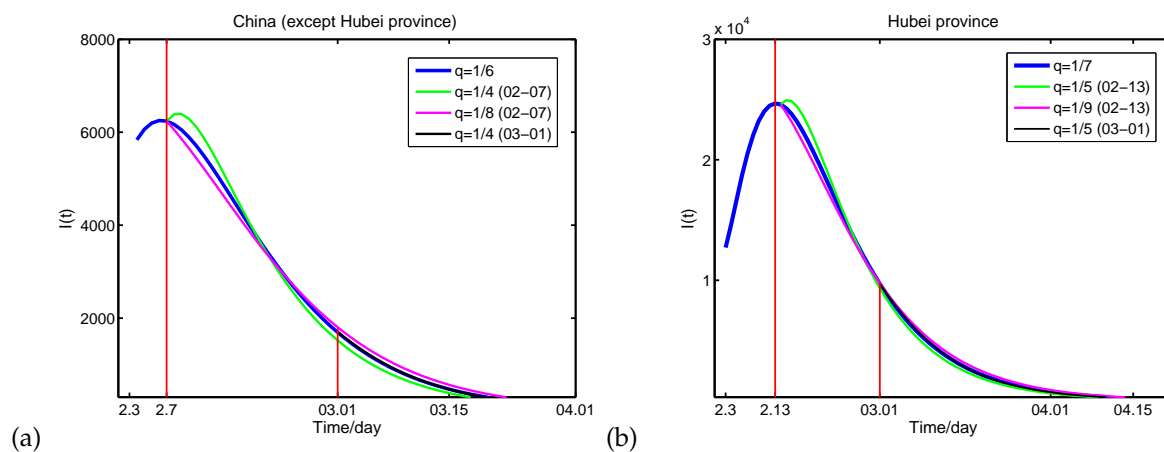


Figure 9. The effects of different strength of detecting efforts on COVID-19 spread in (a) China except Hubei and (b) Hubei province. The red vertical lines mark the peak time of the COVID-19 epidemic that is denoted in blue curves.

predictions for Guangdong, Hunan, Henan, and Jiangxi, as seen in Figure A1. Based our model, the predicted peak arrival time and cumulative confirmed cases are close to the real data of February 12th and the real number of cases of 72,582. Additionally, the current strategies should maintain at least until the middle of March in China, otherwise, it may lengthen the epidemic duration and enlarge the epidemic size (see Figures 8 and 9). Our study has both significant theoretic and practical values. The COVID-19 infections in China will be definitely eliminated under continuous timely and effective public health policies and containment measures.

Author Contributions: conceptualization, X. L., S. F., J. Y., X.-L. P. and Z. J.; methodology, X. L., S. F., J. Y., X.-L. P., H. Z., M.Y. L., H. W. and Z. J.; software, X. L. and J. Y.; validation, J. Z., M. Y., H. Z., M.Y. L. and H. W.; formal analysis, J. Y.; data collection and processing, S. F. and X. C.; writing, X. L., S. F., J. Y. and X.-L. P.; funding acquisition, J. Y., X.-L. P. and Z. J.

Funding: This research was funded by the National Natural Science Foundation of China (grant numbers: 11601294, 61873154, 61573016), and by the Shanxi Research Project on COVID-19 epidemic control and prevention (grant number: 202003D31011/GZ).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

| | |
|----------|--|
| COVID-19 | the novel coronavirus pneumonia |
| MERS | Middle East respiratory syndrome coronarirus |
| NHCC | National Health Commission of China |
| MCMC | Markov Chain Monte Carlo |
| MH | Metropolis-Hasting |

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 for living necessities. 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 [13], the number of family members varies from 1 person to 10 persons. The clustering coefficient ϕ of a network is defined as the number of triangles over 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. Therefore, in this special case of family quarantine, the clustering coefficient is maximal, which takes

$$\phi = \frac{3 \sum_{k=3}^{10} C_k^3 F_k}{\sum_{k=3}^{10} A_k^3 F_k} = \frac{\sum_{k=3}^{10} \binom{k}{3} F_k}{2 \sum_{k=3}^{10} \binom{k}{3} F_k} = 0.5, \quad (\text{A1})$$

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. Ignoring the few addition of triangles, the minimum clustering coefficient in this case is

$$\phi = \frac{3 \sum_{k=3}^{10} C_k^3 F_k}{\sum_{k=3}^{10} A_k^3 F_k + 2(k-1)} = \frac{\sum_{k=3}^{10} \binom{k}{3} F_k}{2 \sum_{k=3}^{10} \binom{k}{3} F_k + 2(k-1)} = 0.42. \quad (\text{A2})$$

It follows from Eqs. (A1) and (A2) 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 $\phi = 0.42$ 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 $\phi = 0.1$. In this paper, we adopted the case of $\phi = 0.1$ 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 latent 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 latent 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 latent 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 latent 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 latent 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.

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 A1 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. It is worth mentioning that we have adopted the reported data of cumulative confirmed cases of COVID-19 infections from January 24th to February 17th, 2020, as the data in these regions are not much influenced by the data reported in Wuhan, the epicenter of the COVID-19 epidemic.

References

1. <http://www.sci-news.com/medicine/2019-ncov-coronavirus-virulence-08068.html>
2. Li Q, Med M, Guan XH, et al., Early transmission dynamics in Wuhan, China, of Novel Coronavirus-infected Pneumonia, *N Engl J Med*, 2020, 10.1056/NEJMoa2001316.
3. Wu P, Hao X, Lau EHY, Wong JY, Leung, K S M, Wu JT, et al. Real-time tentative assessment of the epidemiological characteristics of novel coronavirus infections in Wuhan, China, as at 22 January 2020. *Eurosurveillance*. 2020; 25: 2000044.
4. Health Commission of Hubei Province. (<http://wjw.hubei.gov.cn/>).
5. China News. Mayor of Wuhan: More than 5 million people left Wuhan. January 26, 2020 (<http://ent.chinanews.com/sh/2020/01-26/9070484.shtml>).
6. Tang B, Wang X, Li Q, et al., Estimation of the transmission risk of 2019-nCov and its implication for public health interventions. *J. Clin. Med.*, 2020, 9: 462.
7. Jin GH, Yu JY, Han LY, Duan SW, The impact of traffic isolation in Wuhan on the spread of 2019-nCov, 2020, <https://doi.org/10.1101/2020.02.04.20020438>
8. You C, Deng YH, Hu WJ, et al., Estimation of the time-varying reproduction number of 2019-nCov outbreak in China, 2020, doi: <https://doi.org/10.1101/2020.02.08.20021253>.

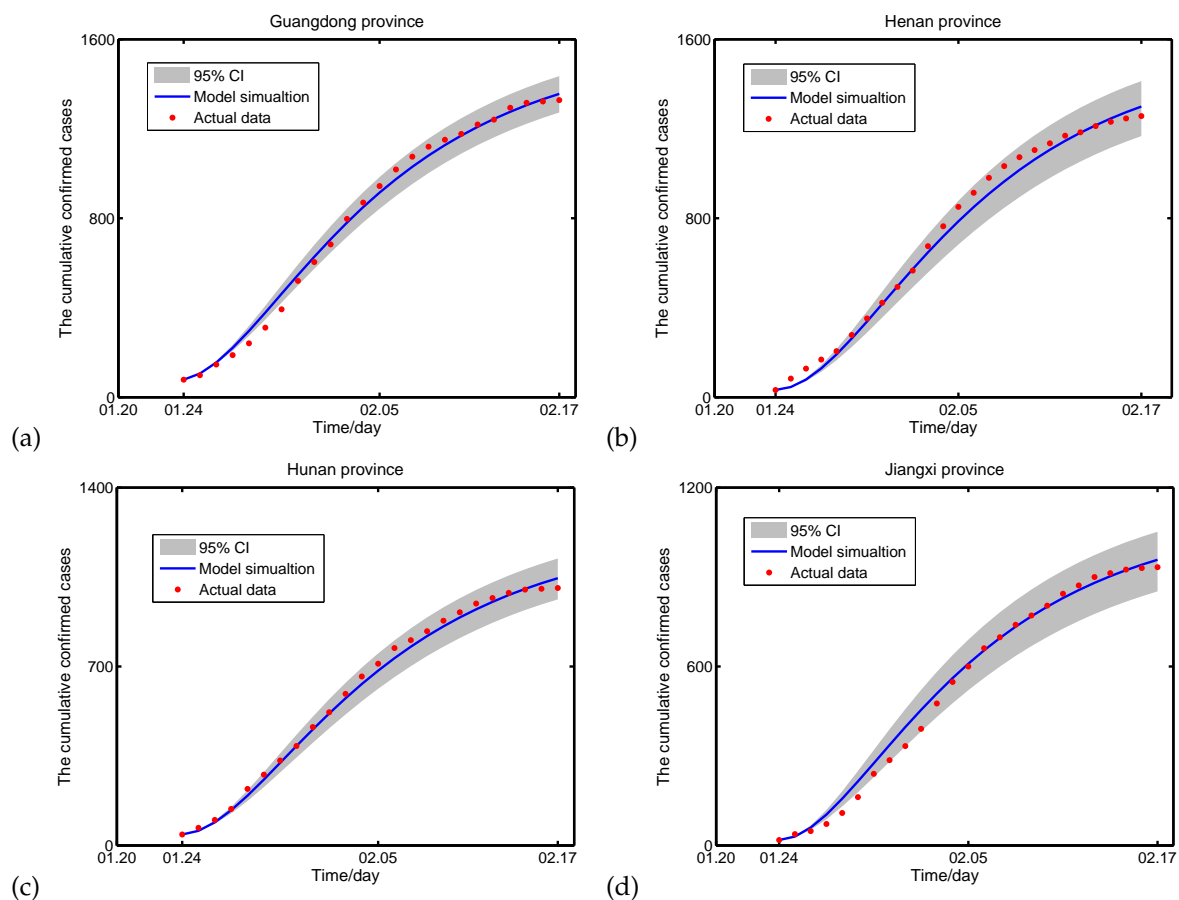


Figure A1. Fitting results of our model to real data of COVID-19 infections in other four hard-hit provinces: (a) Guangdong, (b) Henan, (c) Hunan, (d) Jiangxi.

9. Liu T, Hu JX, Kang M, et al., Transmission dynamics of 2019 novel coronavirus (2019-nCoV), 2020, <https://doi.org/10.1101/2020.01.25.919787>.
10. Chen TM, Rui J, Wang QP, et al., A mathematical model for simulating the transmission of Wuhan novel Coronavirus. 2020, <https://doi.org/10.1101/2020.01.19.911669>.
11. Shao P, Shan YG, Beware of asymptomatic transmission: Study on 2019-nCoV prevention and control measures based on SEIR model, 2020, <https://doi.org/10.1101/2020.01.28.923169>.
12. Shen MW, Peng ZH, Xiao YN, Zhang L, Modelling the epidemic trend of the 2019 novel coronavirus outbreak in China, 2020, <https://doi.org/10.1101/2020.01.23.916726>.
13. 2019 Chinese Statistic Yearbook (in Chinese). <http://www.stats.gov.cn/tjsj/ndsj/2019/indexch.htm>
14. Peng XL, Xu XJ, Small M, Fu XC, Jin Z, Prevention of infectious diseases by public vaccination and individual protection, *J. Math. Biol.*, 2016, 73: 1561.
15. Chen SS, Wang KH, Sun MF, Fu XC, Spread of competing viruses on heterogeneous networks, *Phil. Trans. R. Soc. A*, 2017, 375: 20160248.
16. Thomson NA, Ellner SP. Pari-edge approximation for heterogeneous lattice population models. *Theor. Popul. Biol.*, 2003, 64, 271-280
17. Eames KT. Modelling disease spread through random and regular contacts in clustered populations. *Theor. Popul. Biol.*, 2008, 73: 104-111.
18. Hausken K, Moxnes JF. A closure approximation technique for epidemic models. *Math. Comp. Model. Dyn.* 2010, 16: 555-574.
19. Keeling MJ, Rand DA, Morris AJ, Correlation models for childhood epidemics, *Proc. R. Soc. Lond. B*, 1997, 264: 1149-1156.
20. House T, Keeling MJ, Insights from unifying modern approximations to infections on networks, *J.R. Soc. Interface*, 2011, 8: 67-73.

21. The National Health Commission of the People's Republic of China. <http://www.nhc.gov.cn/>
22. The platform of 2019-nCov infection. <http://www.clas.ac.cn/xwzx2016/163486/xxfysjpt2020/>
23. Official news on the change of diagnosed approaches issued by the Chinese government (in Chinese). http://www.gov.cn/zhengce/zhengceku/2020-02/09/content_5476407.htm
24. Harko T, Lobo FSN, Mak MK, Exact analytical solutions of the Susceptible-infected-recovered epidemic model and of the SIR model with equal death and birth rates, *Appl. Math. Comput.*, 2014, 236: 184-194.
25. Yang J, Xu F. The computational approach for the basic reproduction number of epidemic models on complex networks, *IEEE Access*, 2019, 7: 26474 - 26479.
26. Van Den Driessche, Watmough J, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.*, 2002, 180: 29-48.
27. Bauch CT, Lloyd-Smith JO, Coffee MP, Galvani AP. Dynamically modeling SARS and other newly emerging respiratory illnesses: past, present, and future. *Epidemiology* 2005, 16: 791-801.
28. D. Gamerman, H. F. Lopes, Markov chain Monte Carlo: stochastic simulation for Bayesian inference. Chapman and Hall/CRC, 2006.
29. Wuhan Municipal Health Commission. <http://wjw.wuhan.gov.cn>