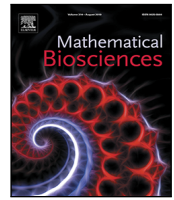




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



## Original Research Article

# A data-driven network model for the emerging COVID-19 epidemics in Wuhan, Toronto and Italy

Ling Xue<sup>a,1</sup>, Shuanglin Jing<sup>a</sup>, Joel C. Miller<sup>b</sup>, Wei Sun<sup>a,1</sup>, Huafeng Li<sup>a</sup>, José Guillermo Estrada-Franco<sup>c,\*</sup>, James M. Hyman<sup>d</sup>, Huaiping Zhu<sup>e,\*</sup>

<sup>a</sup> College of Mathematical Sciences, Harbin Engineering University, Harbin, Heilongjiang, 150001, China

<sup>b</sup> School of Engineering and Mathematical Sciences, Melbourne, La Trobe University, 3086, Australia

<sup>c</sup> Instituto Politécnico Nacional, Centro de Biotecnología Genómica, Cd. Reynosa, Tamaulipas, 88710, México, Mexico

<sup>d</sup> Department of Mathematics, Tulane University, New Orleans, LA, 70118, USA

<sup>e</sup> Lamps and Center of Disease Modelling (CDM), Department of Mathematics and Statistics, York University, Toronto, ON, M3J 1P3, Canada

## ARTICLE INFO

## Keywords:

COVID-19  
Mitigation strategies  
Network model  
Heterogeneity  
Control measures

## ABSTRACT

The ongoing Coronavirus Disease 2019 (COVID-19) pandemic threatens the health of humans and causes great economic losses. Predictive modeling and forecasting the epidemic trends are essential for developing countermeasures to mitigate this pandemic. We develop a network model, where each node represents an individual and the edges represent contacts between individuals where the infection can spread. The individuals are classified based on the number of contacts they have each day (their node degrees) and their infection status. The transmission network model was respectively fitted to the reported data for the COVID-19 epidemic in Wuhan (China), Toronto (Canada), and the Italian Republic using a Markov Chain Monte Carlo (MCMC) optimization algorithm. Our model fits all three regions well with narrow confidence intervals and could be adapted to simulate other megacities or regions. The model projections on the role of containment strategies can help inform public health authorities to plan control measures.

## 1. Introduction

The development of international trade and tourism has accelerated the spatial spread of infectious diseases. The limited data available on emerging epidemics adds to the challenge of mitigating the spread of emerging infections [1]. The unprecedented Coronavirus Disease 2019 (COVID-19) outbreak began at the end of 2019. The number of reported cases keeps rising worldwide and thousands of lives have been claimed. This pandemic is having an enormous impact on world health, disturbing the stability of the societies, and triggers great economic losses. Predicting the future of the pandemic, assessing the impacts of current interventions, and evaluating the effectiveness of alternate mitigation strategies are of utmost importance for saving lives.

Mathematical models can be used to understand the dynamics of epidemics and help inform control strategies. A numerous number of models are being used to project the current COVID-19 pandemic. Ziff and Ziff analyzed the number of reported cases for Wuhan (China) and showed that the growth of the daily number of confirmed new cases indicates an underlying fractal or small-world network of connections between susceptible and infected individuals [2]. Wang et al. developed an SEIR model to estimate the epidemic trends in Wuhan, assuming the

prevention and control measures were either sufficient or insufficient to control the epidemic [3]. Kucharski et al. combined a stochastic transmission model with data on cases of COVID-19 in Wuhan and international cases to estimate how the transmission had varied over time between January and February in 2020 [4].

Kraemer et al. analyzed the impact of interventions on the spread of COVID-19 in China using transportation data [5]. Chinazzi et al. used a global meta-population disease transmission model to project the impact of travel limitations on the national and international spread of the epidemic. They showed that the travel restriction of Wuhan, China had a more marked effect on the international scale than that on Mainland China [6].

Ferguson et al. found that optimal mitigation policies (combining home isolation of suspected cases, home quarantine of those living in the same household as suspected cases, and social distancing of the elderly and others at most risk of severe disease) might reduce peak healthcare demand by 2/3 and deaths by half [7]. Likewise, Hellewell et al. developed a stochastic transmission model and found that highly effective contact tracing and case isolation are enough to control a new outbreak of COVID-19 within three months in most scenarios [8].

\* Corresponding authors.

E-mail addresses: [jestradaf@ipn.mx](mailto:jestradaf@ipn.mx) (J.G. Estrada-Franco), [huaiping@mathstat.yorku.ca](mailto:huaiping@mathstat.yorku.ca) (H. Zhu).

<sup>1</sup> Contributed equally.

Zhang et al. fitted the reported serial interval (mean and standard deviation) with a gamma distribution to estimate the basic reproduction number at the early stage of a COVID-19 outbreak, indicating the potential of second outbreaks [9]. Maier et al. developed a compartmental model dividing individuals into susceptible, exposed, removed, and quarantined symptomatically infected and showed that the distinctive subexponential increase of confirmed cases in mainland China could be explained as a direct consequence of containment policies that effectively deplete the susceptible population [10].

Most of these models are based on assuming the population is homogeneously mixing, that is, the contacts between people are random and uniformly distributed throughout the population. However, different individuals may have varying numbers of acquaintances and contacts in the real world. The important role that heterogenous contact networks play in the transmission dynamics of infectious diseases is often overlooked [11]. Models that take into account contact heterogeneity better represent the actual transmission network through a population and are more likely to capture the true epidemic dynamics.

Disease propagation is closely linked with the structure of social contact networks [12]. The ubiquitous diversity in contact patterns and heterogeneity among individuals depends on differences in social structures, spatial distances, and behavior [13]. The heterogeneity exists at a wide range of scales and leads to highly variable transmission dynamics of infectious diseases [14,15].

Many real-world social networks can be characterized by a random Watts–Strogatz (WS) small-world network [16,17]. In a small-world network, most nodes can be reached from every other node by a small number of hops or steps, even if they are not immediate neighbors. This type of network model allows us to adapt changes to some realistic network structures and examine the effects of control and intervention countermeasures such as social distancing, self-isolation, and personal protection. The framework and analysis can be applied to study the transmission dynamics in different regions and many other infectious diseases.

The COVID-19 epidemic in Wuhan ended in April, while the epidemics in the Greater Toronto Area (GTA, Canada) and the Italian Republic are continuing to grow. We fit the parameters of our network model to the confirmed cases in each of these regions. Although Wuhan, Toronto, and Italy differ in some ways, the way that SARS-CoV-2 is transmitted from one person to another is quite similar. Individuals may acquire infection from other infectious individuals, even if they do not contact each other directly. The Watts–Strogatz model supplies an ideal tool to study the spread of epidemics among individuals even if their locations are not considered. We used the Watts–Strogatz model to generate random networks with the small world properties appropriate for infectious disease transmission in these cities [16,17].

The epidemic curves are all fitted very well using the small-world network structure models, indicating that the typical small-world property is able to capture the contact patterns during COVID-19 epidemics. The differences in these fitted parameters and starting times reflect the differences in the underlying transmission mechanisms and potential spread in the regions. The model then projected the trends of COVID-19 spread by simulating epidemics in the Wuhan, Toronto, and Italy networks. Our findings can guide public health authorities to implement effective mitigation strategies and be prepared for potential future outbreaks.

## 2. The network model

We develop a network-based model by extending the network SIR model [18] by incorporating the characteristics of COVID-19 transmission to assess the spread of the disease in heterogeneous populations. We derive the explicit expression of the epidemic threshold and discuss the final epidemic size for the network model.

### 2.1. Model formulation

We classify individuals by their average number of contacts in a typical day (time unit for the modeling) represented on the network by their degree  $k$  ( $k = 1, 2, \dots, n$ ). Individuals with degree  $k$  are divided into susceptible ( $S_k$ ), exposed ( $E_k$ ), asymptotically infected ( $A_k$ ), symptomatically infected ( $I_k$ ), hospitalized ( $H_k$ ), recovered ( $R_k$ ), and dead ( $D_k$ ) states. Our model is formulated as follows

$$\begin{cases} \frac{ds_k}{dt} &= -\beta k s_k \sum_{k'} [1 - (1 - \frac{k'-1}{k'} P(k'|k) i_{k'}) (1 - \frac{k'-1}{k'} P(k'|k) \sigma a_{k'})], \\ \frac{de_k}{dt} &= \beta k s_k \sum_{k'} [1 - (1 - \frac{k'-1}{k'} P(k'|k) i_{k'}) (1 - \frac{k'-1}{k'} P(k'|k) \sigma a_{k'})] - e e_k, \\ \frac{da_k}{dt} &= (1 - \delta) e e_k - \gamma_a a_k, \\ \frac{di_k}{dt} &= \delta e e_k - \gamma_i i_k - \mu i_k - \xi i_k, \\ \frac{dh_k}{dt} &= \xi i_k - \gamma_h h_k - \mu h_k, \\ \frac{dr_k}{dt} &= \gamma_i i_k + \gamma_a a_k + \gamma_h h_k, \\ \frac{dd_k}{dt} &= \mu i_k + \mu h_k, \end{cases} \quad (2.1)$$

where  $s_k = S_k/N_k$ ,  $e_k = E_k/N_k$ ,  $a_k = A_k/N_k$ ,  $i_k = I_k/N_k$ ,  $h_k = H_k/N_k$ ,  $r_k = R_k/N_k$ , and  $d_k = D_k/N_k$  represent the fractions of susceptible, exposed, asymptotically infected, symptomatically infected, hospitalized, recovered, and dead individuals with degree  $k$ , respectively. Here,  $N_k$  is the total number of individuals with degree  $k$ , and  $N_k = S_k + E_k + A_k + I_k + H_k + R_k + D_k$ , and  $s_k + e_k + a_k + i_k + h_k + r_k + d_k = 1$ .  $P(k'|k)$  represents the probability that an edge from a node with degree  $k$  connects to a node with degree  $k'$ . For uncorrelated networks,  $P(k'|k) = k' P(k')/\langle k \rangle$  [19]. Since the node with degree  $k'$  shares an edge with the node degree  $k$ , and only has  $(k'-1)$  free edges, a fraction  $\frac{k'-1}{k'}$  of nodes may acquire the infection.

We assume that the transmission rates of symptomatically infected individuals and asymptotically infected individuals are  $\beta$  and  $\sigma\beta$ , respectively. The factor  $\sigma$  accounts for the different transmission rates between asymptotically infected individuals and symptomatically infected individuals.  $\beta k s_k \sum_{k'} \frac{k'-1}{k'} P(k'|k) i_{k'}$  represents the fraction of nodes with degree  $k$  infected by symptomatically infected nodes, and  $\sigma \beta k s_k \sum_{k'} \frac{k'-1}{k'} P(k'|k) a_{k'} = \beta k s_k \sum_{k'} \frac{k'-1}{k'} P(k'|k) \sigma a_{k'}$  represents the fraction of nodes with degree  $k$  infected by asymptotically infected nodes. Here,  $\frac{k'-1}{k'} P(k'|k) i_{k'}$  represents the probability that an edge from a degree  $k$  node connects to a symptomatically infected node with degree  $k'$ , and  $\frac{k'-1}{k'} P(k'|k) \sigma a_{k'}$  represents the probability that an edge from a degree  $k$  node connects to an asymptotically infected node with degree  $k'$ .

In Model (2.1), the term  $(1 - \frac{k'-1}{k'} P(k'|k) i_{k'})$  represents the probability of not being infected by a symptomatically infected node with degree  $k'$ , and  $(1 - \frac{k'-1}{k'} P(k'|k) \sigma a_{k'})$  represents the probability of not being infected by an asymptotically infected node with degree  $k'$ . Thus,  $(1 - \frac{k'-1}{k'} P(k'|k) i_{k'}) (1 - \frac{k'-1}{k'} P(k'|k) \sigma a_{k'})$  is the probability that a node will neither be infected by a symptomatically infected nor be infected by an asymptotically infected neighbor with degree  $k'$ , and  $1 - (1 - \frac{k'-1}{k'} P(k'|k) i_{k'}) (1 - \frac{k'-1}{k'} P(k'|k) \sigma a_{k'})$  is the probability of being infected by a symptomatically infected or an asymptotically infected neighbor with degree  $k'$ .

Therefore, the susceptible individuals are infected at rate

$$\beta k s_k \sum_{k'} [1 - (1 - \frac{k'-1}{k'} P(k'|k) i_{k'}) (1 - \frac{k'-1}{k'} P(k'|k) \sigma a_{k'})]$$

and enter the exposed state. After incubation period with a mean time of  $1/\epsilon$  days, exposed individuals become symptomatically infected and asymptotically infected with probabilities  $\delta$  and  $1 - \delta$ , respectively. Symptomatically infected individuals are hospitalized at rate  $\xi$ , and die at rate  $\mu$ . Asymptotically infected individuals, symptomatically infected individuals, and hospitalized individuals recover at rates  $\gamma_a$ ,  $\gamma$ , and  $\gamma_h$ , respectively. Both the hospitalized individuals and symptomatically infected individuals die at rate  $\mu$ .

2.2. Mathematical analysis

We derive the epidemic threshold to predict whether the epidemic will spread or die out and derive final epidemic size to quantify the total number of infected individuals.

2.2.1. The epidemic threshold

To estimate the transmission potential of the epidemic, we derive the important epidemic threshold,  $R_0$ , defined as the average number of secondary cases produced by an infected individual in a completely susceptible population [20]. There exists a disease-free equilibrium,

$$(s_1, \dots, s_n, e_1, \dots, e_n, a_1, \dots, a_n, i_1, \dots, i_n, h_1, \dots, h_n, r_1, \dots, r_n, d_1, \dots, d_n)^T = (1, \dots, 1, 0, \dots, 0, 0, \dots, 0, 0, \dots, 0, 0, \dots, 0, 0, \dots, 0)^T =: P_0.$$

We compute  $R_0$  following the next generation matrix approach presented by van den Driessche and Watmough [21]. For simplicity, we only consider the compartments related to infection, namely,  $e_k$ ,  $a_k$  and  $i_k$ , and rewrite the equations as the difference between vectors  $\mathcal{F}_k$  and  $\mathcal{V}_k$  following the notations in [21]

$$\left[ \frac{de_k}{dt}, \frac{da_k}{dt}, \frac{di_k}{dt} \right]^T = \mathcal{F}_k - \mathcal{V}_k,$$

where

$$\mathcal{F}_k = (F_{ik}) = \begin{bmatrix} \beta k s_k \sum_{k'} [1 - (1 - \frac{k'-1}{k'}) P(k'|k) i_{k'}] (1 - \frac{k'-1}{k'}) P(k'|k) \sigma a_{k'} & & \\ & 0 & \\ & & 0 \end{bmatrix},$$

$$\mathcal{V}_k = (\mathcal{V}_{ik}) = \begin{bmatrix} \epsilon e_k & & \\ -(1 - \delta) \epsilon e_k + \gamma_a a_k & & \\ -\delta \epsilon e_k + \gamma i_k + \mu i_k + \xi i_k & & \end{bmatrix}.$$

Here,  $F_{ik}$  represents the rate at which new infections are produced and  $\mathcal{V}_{ik}$  represents the rate at which individuals transfer between compartments,  $i = 1, 2, 3$  and  $k = 1, \dots, n$  for Model (2.1).

The Jacobian matrix  $F$  is

$$F = \left[ \frac{\partial \mathcal{F}_{ik}}{\partial z_j} \right]_{P_0} = \begin{bmatrix} 0_{n \times n} & \sigma \beta F' & \beta F' \\ 0_{n \times n} & 0_{n \times n} & 0_{n \times n} \\ 0_{n \times n} & 0_{n \times n} & 0_{n \times n} \end{bmatrix},$$

where  $z = (z_j) = (e_1, \dots, e_n, a_1, \dots, a_n, i_1, \dots, i_n)$  and

$$F' = \begin{bmatrix} 0 & \frac{1}{2} P(2|1) & \dots & \frac{n-1}{n} P(n|1) \\ 0 & P(2|2) & \dots & \frac{2(n-1)}{n} P(n|2) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \frac{n}{2} P(2|n) & \dots & (n-1) P(n|n) \end{bmatrix} = \frac{1}{\langle k \rangle} \begin{bmatrix} 0 & P(2) & \dots & (n-1) P(n) \\ 0 & 2P(2) & \dots & 2(n-1) P(n) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & nP(2) & \dots & n(n-1) P(n) \end{bmatrix}.$$

The matrices  $V$  and  $V^{-1}$  are

$$V = \left[ \frac{\partial \mathcal{V}_i}{\partial z_j} \right]_{P_0} = \begin{bmatrix} \epsilon \mathbf{I}_n & 0_{n \times n} & 0_{n \times n} \\ -(1 - \delta) \epsilon \mathbf{I}_n & \gamma_a \mathbf{I}_n & 0_{n \times n} \\ -\delta \epsilon \mathbf{I}_n & 0_{n \times n} & (\gamma + \mu + \xi) \mathbf{I}_n \end{bmatrix},$$

where  $\mathbf{I}_n$  is the  $n \times n$  identity matrix, and

$$V^{-1} = \begin{bmatrix} \frac{1}{\epsilon} \mathbf{I}_n & 0_{n \times n} & 0_{n \times n} \\ \frac{1-\delta}{\gamma_a} \mathbf{I}_n & \frac{1}{\gamma_a} \mathbf{I}_n & 0_{n \times n} \\ \frac{\delta}{\gamma + \mu + \xi} \mathbf{I}_n & 0_{n \times n} & \frac{1}{\gamma + \mu + \xi} \mathbf{I}_n \end{bmatrix}.$$

The next generation matrix is

$$FV^{-1} = \begin{bmatrix} \beta \left( \frac{1-\delta}{\gamma_a} \sigma + \frac{\delta}{\gamma + \mu + \xi} \right) F' & \frac{\sigma \beta}{\gamma_a} F' & \frac{\beta}{\gamma + \mu + \xi} F' \\ 0_{n \times n} & 0_{n \times n} & 0_{n \times n} \\ 0_{n \times n} & 0_{n \times n} & 0_{n \times n} \end{bmatrix}. \tag{2.2}$$

Since the rank of matrix  $F'$  is 1, the spectral radius of  $F'$  is its trace, i.e.,

$$\rho(F') = \text{Tr}(F') = \frac{1}{\langle k \rangle} \sum_k (k-1) k P(k) = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}.$$

It follows from (2.2) that the basic reproduction number  $R_0$  becomes

$$R_0 = \rho(FV^{-1}) = \beta \left( \frac{1-\delta}{\gamma_a} \sigma + \frac{\delta}{\gamma + \mu + \xi} \right) \frac{(\langle k^2 \rangle - \langle k \rangle)}{\langle k \rangle},$$

where  $\beta \frac{1-\delta}{\gamma_a} \sigma$  and  $\beta \frac{\delta}{\gamma + \mu + \xi}$  represent the average numbers of secondary cases produced by an asymptotically infected individual and a symptomatically infected individual in a homogeneously mixed population, respectively. The term  $\frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}$  represents the average excess degree of nodes in the network [22].

2.2.2. Final epidemic size

We shall derive the final size following the approach in [23]. The nonlinear term in the first and second equations of Model (2.1) can be rewritten as

$$\begin{aligned} \beta k s_k \sum_{k'} [1 - (1 - \frac{k'-1}{k'}) P(k'|k) i_{k'}] (1 - \frac{k'-1}{k'}) P(k'|k) \sigma a_{k'} & \\ = \beta k s_k \sum_{k'} [1 - 1 + \frac{k'-1}{k'} P(k'|k) i_{k'} + \frac{k'-1}{k'} P(k'|k) \sigma a_{k'} & \\ - (\frac{k'-1}{k'}) P(k'|k)^2 i_{k'} \sigma a_{k'}] & \\ = \beta k s_k \sum_{k'} [ \frac{k'-1}{k'} P(k'|k) i_{k'} + \frac{k'-1}{k'} P(k'|k) \sigma a_{k'} & \\ - (\frac{k'-1}{k'}) P(k'|k)^2 i_{k'} \sigma a_{k'} ]. & \end{aligned}$$

When  $i_{k'} \ll 1$  and  $a_{k'} \ll 1$ ,  $i_{k'} a_{k'} \approx 0$ . Hence,

$$\begin{aligned} \beta k s_k \sum_{k'} [1 - (1 - \frac{k'-1}{k'}) P(k'|k) i_{k'}] (1 - \frac{k'-1}{k'}) P(k'|k) \sigma a_{k'} & \\ \approx \beta k s_k \sum_{k'} [ \frac{k'-1}{k'} P(k'|k) (i_{k'} + \sigma a_{k'}) ]. & \end{aligned}$$

Hence, Model (2.1) can be simplified as

$$\begin{cases} \frac{ds_k}{dt} = -\beta k s_k \sum_{k'} \frac{k'-1}{k'} P(k'|k) (i_{k'} + \sigma a_{k'}), \\ \frac{de_k}{dt} = \beta k s_k \sum_{k'} \frac{k'-1}{k'} P(k'|k) (i_{k'} + \sigma a_{k'}) - \epsilon e_k, \\ \frac{da_k}{dt} = (1 - \delta) \epsilon e_k - \gamma_a a_k, \\ \frac{di_k}{dt} = \delta \epsilon e_k - \gamma i_k - \mu i_k - \xi i_k, \\ \frac{dh_k}{dt} = \xi i_k - \gamma_h h_k - \mu h_k, \\ \frac{dr_k}{dt} = \gamma i_k + \gamma_a a_k + \gamma_h h_k, \\ \frac{dd_k}{dt} = \mu i_k + \mu h_k. \end{cases} \tag{2.3}$$

We first claim that  $\lim_{t \rightarrow +\infty} e_k(t) = 0$ . Otherwise, there exist  $T, \eta > 0$  such that  $e_k(t) \geq \eta > 0, \forall t > T$  since  $e_k(t) \geq 0$ . Since  $\frac{d}{dt}(s_k + e_k) = -\epsilon e_k \leq -\epsilon \eta, s_k(t) + e_k(t) \leq -\epsilon \eta t + s(T) + e(T)$ . Therefore,  $\lim_{t \rightarrow \infty} (s_k(t) + e_k(t)) = -\infty$ , leading to a contradiction. Similarly, we can show that

$$\lim_{t \rightarrow +\infty} e_k(t) = \lim_{t \rightarrow +\infty} a_k(t) = \lim_{t \rightarrow +\infty} i_k(t) = \lim_{t \rightarrow +\infty} h_k(t) = 0, \quad \forall k.$$

For a homogeneous network where all nodes have identical degree  $k$ , Model (2.3) can be reduced to the following model

$$\begin{cases} \frac{ds}{dt} = -\beta k s \left( \frac{k-1}{k} \right) (i + \sigma a), \\ \frac{de}{dt} = \beta k s \left( \frac{k-1}{k} \right) (i + \sigma a) - \epsilon e, \\ \frac{da}{dt} = (1 - \delta) \epsilon e - \gamma_a a, \\ \frac{di}{dt} = \delta \epsilon e - (\gamma + \mu + \xi) i, \\ \frac{dh}{dt} = \xi i - \gamma_h h - \mu h, \\ \frac{dr}{dt} = \gamma i + \gamma_a a + \gamma_h h, \\ \frac{dd}{dt} = \mu i + \mu h, \end{cases} \tag{2.4}$$

with the initial values  $s(0) = s_0$ ,  $e(0) = e_0$ ,  $a(0) = a_0$ ,  $i(0) = i_0$ ,  $h(0) = h_0$ ,  $r(0) = 0$  and  $d(0) = 0$ .

By Model (2.4) and a direct calculation, we have

$$-(i + \sigma a) = \frac{1}{\gamma + \mu + \xi} \left( \frac{ds}{dt} + \frac{de}{dt} + \frac{da}{dt} + \frac{di}{dt} \right) + \left( \frac{\sigma}{\gamma_a} - \frac{1}{\gamma + \mu + \xi} \right) \left( (1 - \delta) \left( \frac{ds}{dt} + \frac{de}{dt} \right) + \frac{da}{dt} \right). \tag{2.5}$$

By the first equation in Model (2.4), we further have

$$\ln \frac{s(+\infty)}{s_0} = -\beta k \left\langle \frac{k-1}{k} \right\rangle \int_0^{+\infty} (i + \sigma a) dt, \tag{2.6}$$

where  $s(+\infty) = \lim_{t \rightarrow \infty} s(t)$ . To determine the final size of susceptible individuals,  $s(+\infty)$ , we set

$$f(x) = s_0 \exp \left\{ \beta k \left\langle \frac{k-1}{k} \right\rangle \left[ \frac{x - y_0 - a_0 - i_0}{\gamma + \mu + \xi} + \left( \frac{\sigma}{\gamma_a} - \frac{1}{\gamma + \mu + \xi} \right) \left( (1 - \delta) (x - y_0) - a_0 \right) \right] \right\},$$

where  $y_0 = s_0 + e_0$ . By (2.5), (2.6) and the definition of  $f(x)$ , we have  $s(+\infty) = f(s(+\infty))$ .

It is clear that  $f(x)$  is a positive, increasing, strictly convex function, and  $f(s_0) < s_0$ . Thus,  $f$  has a unique fixed point  $s^+$  in the interval  $(0, s_0)$ , which can be calculated numerically by using the iteration method and

$$s^+ = \lim_{m \rightarrow +\infty} f^m(s_0),$$

where  $f^m$  denotes composition of  $f$  for  $m$  times. Then, the final size of susceptible individuals for a homogeneous network,  $s(+\infty)$ , can be determined by  $s^+$ .

We now derive the final size for heterogeneous networks. Integrating the first equation in Model (2.3) from 0 to  $t$ , we have

$$\ln \frac{s_k(t)}{s_k(0)} = -\beta k \sum_{k'} \frac{k' - 1}{k'} P(k'|k) \int_0^t (i_{k'} + \sigma a_{k'}) du. \tag{2.7}$$

By summing and integrating the equations in Model (2.3),

$$\int_0^t i_k du = -\frac{\delta}{\gamma + \mu + \xi} (y_k(t) - y_k(0)) - \frac{i_k(t) - i_k(0)}{\gamma + \mu + \xi}, \tag{2.8}$$

and

$$\int_0^t a_k du = -\frac{1 - \delta}{\gamma_a} (y_k(t) - y_k(0)) - \frac{1}{\gamma_a} (a_k(t) - a_k(0)), \tag{2.9}$$

where  $y_k(t) = s_k(t) + e_k(t)$ . We set

$$g_k(t) = \left( \frac{\delta}{\gamma + \mu + \xi} + \sigma \frac{1 - \delta}{\gamma_a} \right) y_k(t) + \frac{\sigma}{\gamma_a} a_k(t) + \frac{i_k(t)}{\gamma + \mu + \xi}.$$

By Eqs. (2.7), (2.8) and (2.9), we have

$$\begin{aligned} \ln \frac{s_k(+\infty)}{s_k(0)} &= \beta k \sum_{j=1}^n \frac{j-1}{j} P(j|k) (g_j(+\infty) - g_j(0)) \\ &= \beta k \sum_{j=1}^n \frac{j-1}{j} P(j|k) \left( \left( \frac{\delta}{\gamma + \mu + \xi} + \sigma \frac{1 - \delta}{\gamma_a} \right) s_j(+\infty) - g_j(0) \right), \end{aligned}$$

where  $g_k(0) \geq 0, \forall k$ . Therefore, for all  $k = 1, \dots, n$ , the final size of susceptible individuals satisfies

$$s_k(+\infty) = s_k(0) \exp \left\{ \beta k \sum_{j=1}^n \frac{j-1}{j} P(j|k) \cdot \left( \left( \frac{\delta}{\gamma + \mu + \xi} + \sigma \frac{1 - \delta}{\gamma_a} \right) (s_j(+\infty) - s_j(0)) - w_j(0) \right) \right\},$$

where

$$w_k(0) = \left( \frac{\delta}{\gamma + \mu + \xi} + \sigma \frac{1 - \delta}{\gamma_a} \right) e_k(0) + \frac{\sigma}{\gamma_a} a_k(0) + \frac{i_k(0)}{\gamma + \mu + \xi} \geq 0.$$

We define a map  $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $x = (x_j) \mapsto G(x) = (G_1(x), \dots, G_n(x))^T$  by

$$G_i(x) = s_i(0) \exp \left\{ \beta i \sum_{j=1}^n \frac{j-1}{j} P(j|i) \cdot \left( \left( \frac{\delta}{\gamma + \mu + \xi} + \sigma \frac{1 - \delta}{\gamma_a} \right) (x_j - s_j(0)) - w_j(0) \right) \right\}.$$

To analyze the properties of  $G(x)$ , we shall introduce some notations. For  $Y = (Y_1, \dots, Y_n)^T$ ,  $Z = (Z_1, \dots, Z_n)^T \in \mathbb{R}^n$ , we denote

$$Y \leq Z \text{ (resp. } Y \ll Z) \text{ if } Y_l \leq Z_l \text{ (resp. } Y_l < Z_l), \forall l = 1, \dots, n. \tag{2.10}$$

Moreover, we shall claim  $Y < Z$  if  $Y \leq Z$  and  $Y \neq Z$ . The above definition defines a partial order in  $\mathbb{R}^n$ . For later use, we could extend this partial order to  $n \times n$  matrices as follows. For any  $n \times n$  matrices  $A, B$ , we have

$$A \leq B \text{ if } Ax \leq Bx, \forall 0 \leq x \in \mathbb{R}^n.$$

When  $0 \ll s(0) = [s_1(0), \dots, s_n(0)]^T$  and  $0 \leq w(0) = [w_1(0), \dots, w_n(0)]^T$ , by the definition of  $G(x)$  and partial order defined in (2.10), we have

$$0 \ll G(0) \leq G(s(0)) \leq s(0),$$

Since each  $G_i(x)$  is a increasing function, we have

$$0 \ll G(0) \leq \dots \leq G^m(0) \leq G^m(s(0)) \leq \dots \leq G(s(0)) \leq s(0),$$

where  $G^m$  is the composition function of  $G$  for  $m$  times. By the monotone criterion, we obtain

$$0 \ll \underline{s} := \lim_{m \rightarrow +\infty} G^m(0) \leq \bar{s} := \lim_{m \rightarrow +\infty} G^m(s(0)) \leq s(0).$$

Due to the continuity of  $G$ ,  $G(\underline{s}) = \underline{s}$  and  $G(\bar{s}) = \bar{s}$ . Therefore, we have the following property [23].

**Lemma 2.1.** *All the fixed points of  $G$  in the interval  $[0, s(0)]$  are contained in  $[\underline{s}, \bar{s}]$ .*

Due to the continuous differentiability of  $G$ ,

$$\frac{\partial G_i(x)}{\partial x_j} = \beta i \frac{j-1}{j} P(j|i) \left( \frac{\delta}{\gamma + \mu + \xi} + \sigma \frac{1 - \delta}{\gamma_a} \right) G_i(x) \tag{2.11}$$

for any  $x \in \mathbb{R}^n$  and  $1 \leq i, j \leq n$ . Moreover, we shall simply write (2.11) in terms of the matrix form by

$$DG(x) := \frac{\partial G(x)}{\partial x} = \text{diag} (G_1(x), \dots, G_n(x)) B,$$

where  $B = [b_{ij}]$  and  $b_{ij} = \beta i \frac{j-1}{j} \left( \frac{\delta}{\gamma + \mu + \xi} + \sigma \frac{1 - \delta}{\gamma_a} \right) P(j|i)$ .

By the monotony of  $G$ ,  $DG$  is also monotonous, i.e.,  $DG(x) \leq DG(y)$  for any  $0 \leq x \leq y \leq s(0)$ . By utilizing the properties of  $w(x)$  and  $G(x)$ , we can obtain the following theorem.

**Theorem 2.1.** *Assume that the network is connected, we have*

- (1)  $w(0) = 0$  if and only if  $G(s(0)) = s(0)$ ;
- (2) when  $w(0) > 0$ ,  $G$  has a unique fixed point  $s^{++}$  satisfying  $0 \ll s^{++} < s(0)$ .

The proof of Theorem 2.1 directly follows the proof in [23]. Hence, the final size of susceptible individuals for a heterogeneous network,  $s(+\infty)$ , can be determined by  $s^{++}$  to quantify the number of susceptible individuals left theoretically.

### 3. Parameter estimation and model based forecasting

We parameterized the model with reported data on COVID-19 cases and presented forecasts of the epidemic trends for the three areas.

**Table 1**  
The parameter values and initial condition of four-phase simulations for Wuhan.

Parameter	Mean value	Std	95% CI	References
$\beta$ (Phase 1)	0.04644	$2.83 \times 10^{-3}$	[0.0409, 0.0520]	MCMC
$\beta$ (Phase 2)	0.01597	$4.79 \times 10^{-3}$	$[6.59 \times 10^{-3}, 0.0254]$	MCMC
$\beta$ (Phase 3)	$2.90 \times 10^{-4}$	$2.24 \times 10^{-4}$	$[0, 7.29 \times 10^{-4}]$	MCMC
$\beta$ (Phase 4)	$7.30 \times 10^{-5}$	$7.27 \times 10^{-5}$	$[0, 2.16 \times 10^{-4}]$	MCMC
$\xi$ (Phase 1)	0.8550	0.1220	[0.6159, 1.0942]	MCMC
$\xi$ (Phase 2)	0.1499	0.1557	[0, 0.4552]	MCMC
$\xi$ (Phase 3)	0.3369	0.1089	[0.1234, 0.5504]	MCMC
$\xi$ (Phase 4)	0.8764	0.0836	[0.7124, 1.0403]	MCMC
$\sum_k N_k(0)$	11081000	–	–	[27]
$\sum_k S_k(0)$	11080770	–	–	Calculated
$\sum_k E_k(0)$	200.26	18.88	[163.25, 237.26]	MCMC
$\sum_k A_k(0)$	17.81	4.13	[9.71, 25.90]	MCMC
$\sum_k I_k(0)$	11.78	6.06	[0, 23.66]	MCMC
$\sum_k H_k(0)$	41	–	–	[25]
$\sum_k R_k(0)$	0	–	–	Estimated
$\sum_k D_k(0)$	0	–	–	Estimated

Phase 1 is between Jan 11 and Jan 23, Phase 2 is between Jan 23 and Feb 1, Phase 3 is between Feb 1 and Feb 12, and Phase 4 is between Feb 12 and Mar 31.

### 3.1. Fitting reported confirmed cases

We simulated the spread of COVID-19 in Wuhan, Toronto, and Italy on the Watts–Strogatz network with degree  $k_{min} = 1$  and  $k_{max} = 10$ .

The study period for Wuhan starts from January 11, 2020, after the confirmed cases were reported, the public becomes aware of the infection and most people are trying to avoid gathering. The study period starts from January 26 for Toronto and from January 31 for Italy. In Toronto and Italy, usually people do not gather, especially after lockdown on Wuhan city, the awareness of avoiding exposure to the virus is increasing. Most people stay home during the study period, and the family sizes in Wuhan, Toronto, and Italy on average are all around 3. Therefore, the range of the node degrees is assumed to be between 1 and 10.

The Watts–Strogatz model starts with a ring of  $N$  vertices in which each vertex is connected to its  $2m$  nearest neighbors ( $m$  vertices clockwise and  $m$  counterclockwise). Each edge is connected to a clockwise neighbor with probability  $p$  and preserved with probability  $1 - p$  [19], where the degree distribution is

$$P(k) = \sum_{n=0}^{\min(k-m,m)} \binom{m}{n} (1-p)^n p^{m-n} \frac{(pm)^{k-m-n}}{(k-m-n)!} e^{-pm}.$$

When  $p \rightarrow 1$ , the expression reduces to a Poisson distribution as follows

$$P(k) = \frac{m^{k-m}}{(k-m)!} e^{-m}.$$

In the simulations, we used this degree distribution.

The total number of nodes for Wuhan, Toronto, and Italy are 11081000, 5928000, and 59430000 as shown in Table 1, Table 3, and Table 6, respectively. We parameterized the model using the MCMC approach [24] by MATLAB R2016a according to the number of newly confirmed cases and the cumulative number of cases reported by the Health Commission of Hubei Province [25] and WHO [26].

The rate at which the fraction of the cumulative number of cases changes is  $dc_k/dt = \xi_i k$ , where  $c_k(t)$  represents the fraction of the cumulative number of infected individuals with degree  $k$ . The number of newly infected can be expressed as

$$P_k = [c_k(t) - c_{k-1}(t)]N_k,$$

where  $P_k$  represents the number of new cases with degree  $k$ , and  $N_k$  represents the total number of individuals with degree  $k$ . We run the MCMC simulation for 20000 iterations to fit the value of  $P_k$ .

Zhou et al. showed that the median time from illness onset (i.e., before admission) to discharge was 22 days (IQR 18–25), whereas the median time to death was 18.5 days with IQR between 15 and 22 days [28]. We assume an exponential distribution for the time to

**Table 2**  
Basic reproduction numbers computed by MCMC on the WS network.

Location	Period	Mean value	Standard derivation	95% CI
Wuhan	Jan 11–Jan 23	3.4074	0.2099	[2.9959, 3.8188]
	Jan 23–Feb 1	1.3065	0.3976	[0.5273, 2.0858]
	Feb 1–Feb 12	0.0221	0.0170	[0, 0.0555]
	Feb 12–Mar 31	$5.35 \times 10^{-3}$	$5.34 \times 10^{-3}$	[0, 0.0158]
Toronto	Jan 26–Mar 18	0.6416	0.0867	[0.4716, 0.8116]
	Mar 18–Mar 29	0.0115	0.0151	[0, 0.0412]
Italy	Jan 31–Mar 8	1.4763	0.0984	[1.2834, 1.6691]
	Mar 8–Mar 26	0.0359	0.0185	[0, 0.0721]

recovery for asymptotically infected individuals, symptomatically infected individuals, and hospitalized individuals. This results in the recovery rates  $\gamma_a = \gamma = \gamma_h = 1/22$  per day, and the mortality rate,  $\mu$  is  $1/18.5$  per day. The incubation period of COVID-19 is around 7 days [4], resulting in the progression rate  $\epsilon = 1/7$ . Qiu et al. reported that around 30%–60% of people infected with COVID-19 are asymptomatic or only have mild symptoms, and their transmissibility is lower, but still significant [29]. Thus, we assume that the probability that an infected individual is asymptomatic is  $1 - \delta = 0.6$ , and  $\sigma = 1$  for simulations.

We divided the Wuhan epidemic into four phases according to the reported data [3]. The first phase is before lockdown on Jan 23, 2020. The second phase is between Jan 24, 2020 and Feb 1, 2020 when the hospitals were short of beds. The third phase is between Feb 2, 2020 and Feb 6, 2020 when the Thunder God Mountain Hospital (TGMH) and Fire God Mountain Hospital (FGMH) were put into use. The fourth phase began when door-to-door screening was implemented on Feb 7, 2020 and TGMH, FGMH, and Mobile Cabin Hospitals (MCH) were put into use.

The study period for Toronto (Canada) was decomposed into two phases, namely, the period before Mar 18 and the period after Mar 18 when the city announced the emergence and schools and universities in Toronto were closed on Mar 18.

The study period for Italy was divided into two phases. The early epidemic phase was between Jan 31, 2020 and Mar 8, 2020 when the infection was spreading through the northern provinces. The second period begins on Mar 9, 2020 when the national lockdown started.

### 3.2. Predicting future epidemic trends

The parameters and initial conditions of simulations for Wuhan on the WS network are shown in Table 1. The probability of transmission through adequate contact is estimated by MCMC. The 5000 realizations of the basic reproduction numbers derived for Wuhan using the parameter values listed in Table 1 are shown in Table 2.

From Jan 11 to Mar 31, we estimate that the mean reproduction number on the WS network decreases from 3.41 in the first phase to  $5.35 \times 10^{-3}$  in the fourth phase. The epidemic on the WS network is shown in Fig. 1. Up to Jan 23, 2020 when Wuhan lockdown started, the estimated epidemic size is  $3.96 \times 10^6$ . During the second stage, after the lockdown of Wuhan and before the TGMH and FGMH were put into use, the predicted final size is  $2.17 \times 10^6$ . Thus, the lockdown of Wuhan reduced the expected final size by 45.22%. During the third stage, after TGMH and FGMH were put into use, the final size is  $1.02 \times 10^5$ . Hence, the city lockdown and the usage of TGMH and FGMH reduced the final size by 97.42%. During the fourth stage, after MCH was put into use, the predicted final size is 51269, and the expected final size of infection is reduced by 98.70% due to the increase of healthcare capacity.

The variability of the numbers of confirmed new cases is consistent with the variability of the reproduction numbers listed in Table 2. In the first two phases, the epidemic spread rapidly with larger reproduction numbers that are larger than one, and the numbers of

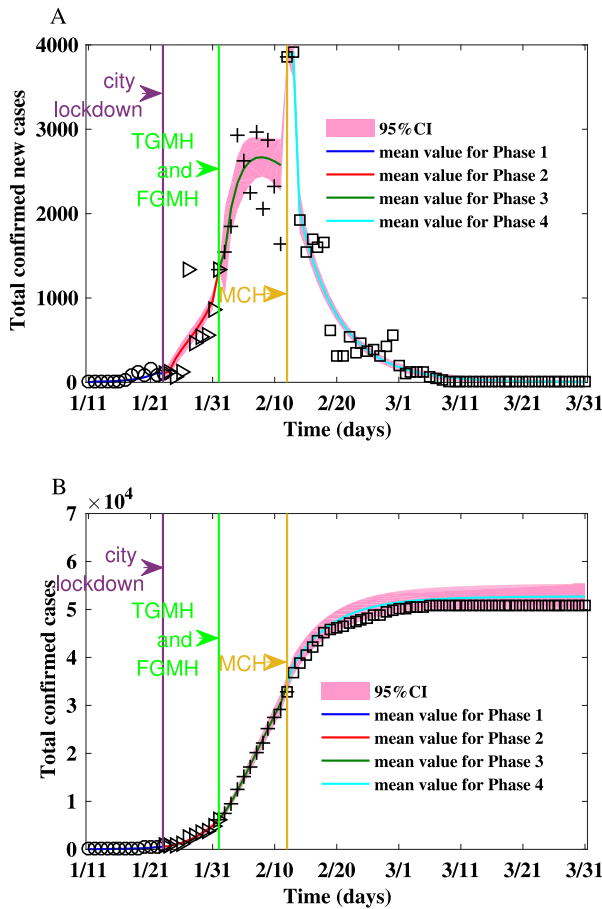


Fig. 1. Fitting the number of reported new cases and the cumulative number of reported cases between Jan 11, 2020 and Mar 31, 2020 for Wuhan on Watts–Strogatz network. (A) Fitting the number of reported new cases on the Watts–Strogatz network. (B) Fitting the cumulative number of reported cases on the Watts–Strogatz network.

Table 3

The parameter values and initial condition of simulations for Toronto.

Parameter	Mean value	Std	95% CI	References
$\beta$ (Phase 1)	$7.95 \times 10^{-3}$	$1.12 \times 10^{-3}$	$[5.76 \times 10^{-3}, 0.01]$	MCMC
$\beta$ (Phase 2)	$1.38 \times 10^{-4}$	$1.79 \times 10^{-4}$	$[0, 4.90 \times 10^{-4}]$	MCMC
$\xi$ (Phase 1)	0.1421	0.0734	$[0, 0.2858]$	MCMC
$\xi$ (Phase 2)	0.1140	0.0709	$[0, 0.2530]$	MCMC
$\sum_k N_k(0)$	5928000	-	-	[30]
$\sum_k S_k(0)$	5927990	-	-	Calculated
$\sum_k E_k(0)$	4.12	4.83	$[0, 13.58]$	MCMC
$\sum_k A_k(0)$	2.43	2.22	$[0, 6.78]$	MCMC
$\sum_k I_k(0)$	3.22	2.89	$[0, 8.88]$	MCMC
$\sum_k H_k(0)$	1	-	-	[26]
$\sum_k R_k(0)$	0	-	-	Estimated
$\sum_k D_k(0)$	0	-	-	Estimated

Phase 1 is from Jan 26 to Mar 18, and Phase 2 is from Mar 18 to Mar 29.

infected cases increase. In the last two phases, the spread is contained as the reproduction numbers falling below 1. In the third phase,

Table 4

The peak number of new cases, peak time, and final epidemic size after containment strategies are implemented in Toronto.

Scenarios	Peak size (95% CI)	Peak time (95% CI)	Final size (95% CI)
$\beta$	60.19 (47.42, 72.97)	Apr 2 (Mar 29, Apr 7)	2712 (1603, 3820)
$0.8\beta$	50.30 (41.03, 59.57)	Apr 1 (Mar 28, Apr 6)	2217 (1451, 2984)
$0.6\beta$	40.94 (34.55, 47.32)	Mar 31 (Mar 26, Apr 4)	1751 (1239, 2262)
$q = 0$	60.19 (47.42, 72.97)	Apr 2 (Mar 29, Apr 7)	2712 (1603, 3820)
$q = 1/8$	44.66 (39.12, 50.21)	Mar 28 (Mar 26, Mar 31)	1486 (1111, 1861)
$q = 1/4$	38.79 (33.86, 43.72)	Mar 27 (Mar 25, Mar 29)	1133 (878, 1388)

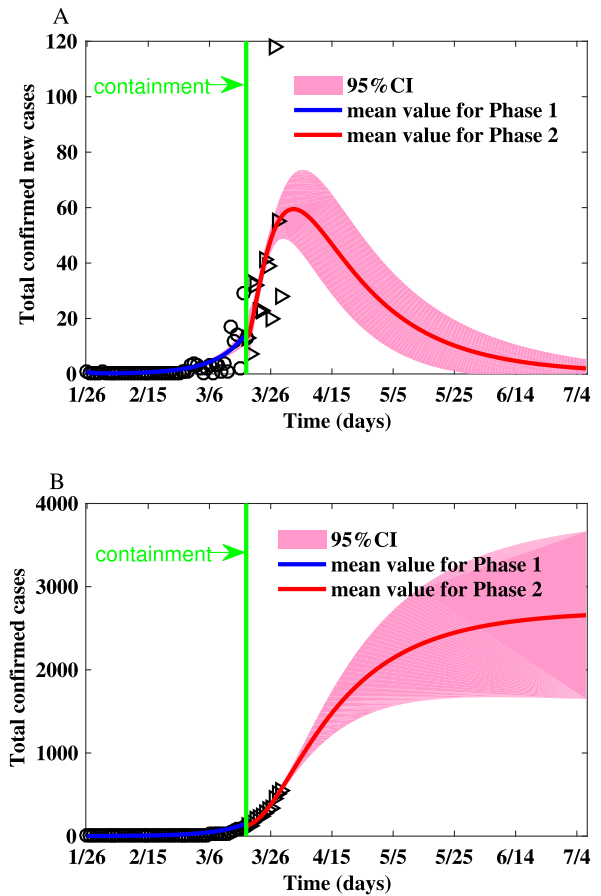


Fig. 2. Fitting the number of reported new cases and the cumulative number of reported cases for Toronto on the Watts–Strogatz network. (A) Fitting the number of reported new cases on the Watts–Strogatz network. (B) Fitting the cumulative number of reported cases on the Watts–Strogatz network.

because a large number of cases are confirmed by door-to-door screening and expanded healthcare capacity, the cumulative number of confirmed cases increased. On the other hand, the epidemic will die out because the reproduction number is less than one. In the fourth phase, the epidemic has been under control with the reproduction number being less than one. Hence, the number of new cases decreases.

The parameters and initial conditions of simulations for the GTA are shown in Table 3. The 5000 realizations of the basic reproduction numbers derived for Toronto using the parameter values listed in Table 3 are shown in Table 2. The reproduction numbers are much smaller due to social distancing policy, school closure, as well as behavior changes. The summary of the simulations is shown in Tables 4 and 5 and simulation results are shown in Fig. 2. The peak size is 60.19 (95% CI: 47.42–72.97), the peak time is Apr 2 (95% CI: Mar 29–Apr 7), and the final size is 2712 (95% CI: 1603–3820).

The parameters and initial condition of simulations for Italy is shown in Table 6. The 5000 realizations of the basic reproduction numbers derived for Italy using the parameter values listed in Table 6

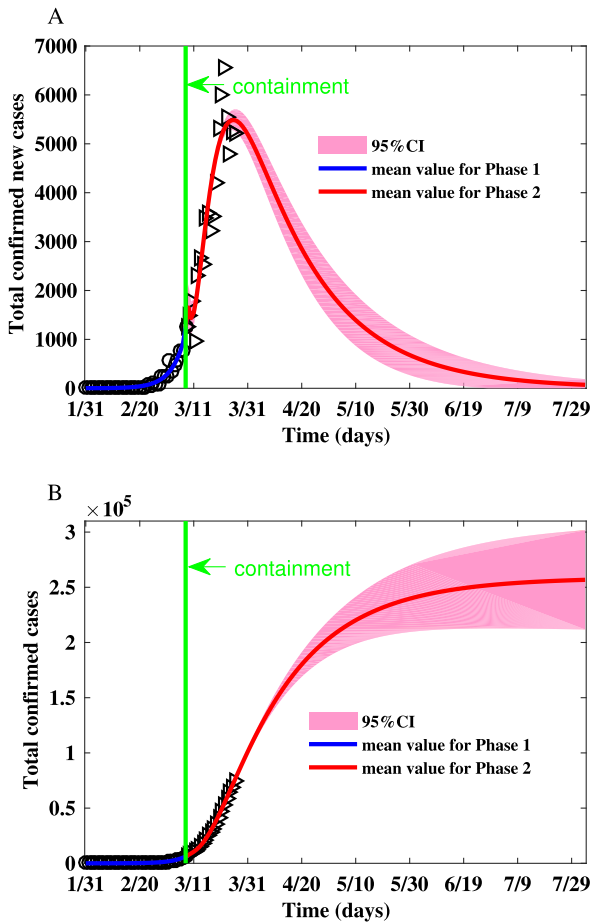


Fig. 3. Fitting the number of reported new cases and the cumulative number of reported cases for Italy on the Watts–Strogatz network. (A) Fitting the number of reported new cases on the Watts–Strogatz network. (B) Fitting the cumulative number of reported cases on the Watts–Strogatz network.

Table 5

The peak number of new cases, peak time, and final epidemic size for Toronto when varying node degrees of symptomatically infected individuals.

Degree	Peak size (95% CI)	Peak time (95% CI)	Final size (95% CI)
$k$	60.19 (47.42, 72.97)	Apr 2 (Mar 29, Apr 7)	2712 (1603, 3820)
$k - 1$	51.92 (42.08, 61.77)	Apr 1 (Mar 28, Apr 6)	2301 (1476, 3125)
$k - 2$	43.98 (36.76, 51.20)	Mar 31 (Mar 27, Apr 5)	1908 (1310, 2506)
$k - 3$	36.61 (30.96, 42.26)	Mar 30 (Mar 25, Apr 3)	1531 (1114, 1949)

Table 6

The parameter values and initial condition of simulations for Italy.

Parameter	Mean value	Standard derivation	95% CI	References
$\beta$ (Phase 1)	0.0179	$1.19 \times 10^{-3}$	[0.0156, 0.0203]	MCMC
$\beta$ (Phase 2)	$4.45 \times 10^{-4}$	$2.30 \times 10^{-4}$	$[0.895 \times 10^{-4}]$	MCMC
$\xi$ (Phase 1)	0.0996	0.0458	$[9.87 \times 10^{-3}, 0.1894]$	MCMC
$\xi$ (Phase 2)	0.1312	0.0250	[0.0823, 0.1801]	MCMC
$\sum_k N_k(0)$	59430000	–	–	[31]
$\sum_k S_k(0)$	59429892	–	–	Calculated
$\sum_k E_k(0)$	69.01	56.92	[0, 180.58]	MCMC
$\sum_k A_k(0)$	22.60	18.24	[0, 58.35]	MCMC
$\sum_k I_k(0)$	32.11	24.23	[0, 79.59]	MCMC
$\sum_k H_k(0)$	2	–	–	[26]
$\sum_k R_k(0)$	0	–	–	Estimated
$\sum_k D_k(0)$	0	–	–	Estimated

Phase 1 is between Jan 31 and Mar 8, and Phase 2 is between Mar 8 and Mar 26.

are shown in Table 2. The reproduction numbers in the second phase are much smaller than that in the first phase due to the awareness of

the severity of the epidemic. The summary of the simulation results is shown in Tables 7 and 8. Fig. 3 shows that the peak number of new cases is 5492 (95% CI: 5277–5708) on Mar 26 (95% CI: Mar 24–Mar 27), and the final size is  $2.59 \times 10^5$  (95% CI:  $2.10 \times 10^5 - 3.08 \times 10^5$ ).

#### 4. The impact of mitigation strategies

The close contacts identified by contact tracing will be quarantined due to exposure to COVID-19 to see if they become sick. To evaluate the impact of mitigation strategies on the spread of COVID-19, Model (2.1) is rewritten as follows

$$\begin{cases}
 \frac{ds_k}{dt} = -\beta k s_k \sum_{k'} [1 - (1 - \frac{k'-1}{k'} P(k'|k) i_{k'}) (1 - \frac{k'-1}{k'} P(k'|k) \sigma a_{k'})] - q s_k + \lambda s q_k, \\
 \frac{dsq_k}{dt} = q s_k - \lambda s q_k, \\
 \frac{de_k}{dt} = \beta k s_k \sum_{k'} [1 - (1 - \frac{k'-1}{k'} P(k'|k) i_{k'}) (1 - \frac{k'-1}{k'} P(k'|k) \sigma a_{k'})] - e e_k, \\
 \frac{da_k}{dt} = (1 - \delta) e e_k - \gamma_a a_k, \\
 \frac{di_k}{dt} = \delta e e_k - \gamma_i i_k - \mu i_k - \xi i_k, \\
 \frac{dh_k}{dt} = \xi i_k - \gamma_h h_k - \mu h_k, \\
 \frac{dr_k}{dt} = \gamma_i i_k + \gamma_a a_k + \gamma_h h_k, \\
 \frac{dd_k}{dt} = \mu i_k + \mu h_k,
 \end{cases}
 \tag{4.1}$$

where  $s q_k = S Q_k / N_k$  represents the fraction of quarantined individuals with degree  $k$ . The parameter  $q$  represents the rate at which susceptible individuals are quarantined, and  $\lambda$  represents the rate at which the quarantined and uninfected close contacts transfer to the susceptible compartment again. In the simulations, we let  $\lambda = 1/14$  to approximate a mean time of 14 days in the exposed state.

For Wuhan, the cumulative number of infected individuals after lockdown and TGMH, FGMH, as well as MCH were put into use are shown in Fig. 4. The results show that the lockdown and the increase in healthcare capacity are effective in controlling the numbers of confirmed cases.

For Toronto, the number of newly infected individuals and the cumulative number of infected individuals produced on the WS network after implementing additional containment strategies besides school closure are shown in Fig. 5. We simulated the scenarios of implementing various containment strategies for Toronto. Simulation results showed that personal protection, reducing the node degrees of symptomatically infected individuals, and quarantine of close contacts are effective in reducing the peak epidemic size and final epidemic size. When  $\beta$  is reduced by 20% by personal protection or social distancing, the peak occurs one day earlier, and the final epidemic size is reduced by around 18%. When  $\beta$  is reduced by 40%, the peak occurs two days earlier, and the final epidemic size is reduced by around 33.3%. When  $q = 1/8$ , the peak occurs four days earlier, and the final epidemic size is reduced by 45.21%. When  $q = 1/4$ , the peak appears five days earlier, and the final epidemic size is reduced by 58.22%. When the node degrees of symptomatically infected individuals are reduced by 1, 2, and 3, the number of new cases produced per day at the peak is reduced by 13.74%, 26.93%, and 39.18%. The final epidemic size is reduced by 15.15%, 29.65%, and 43.55% when the node degrees of symptomatically infected individuals are reduced by 1, 2, and 3, respectively.

For Italy, the number of newly infected individuals and the cumulative number of infected individuals simulated on the WS network after implementing hypothetical containment strategies are shown in Fig. 6. Various scenarios of implementing mitigation strategies showed that the peak epidemic size and final epidemic size in Italy are greatly reduced by personal protection, social distancing, behavior change of symptomatically infected individuals, and quarantine. The simulations



**Table 7**

The peak number of new cases, peak time, and final size after containment strategies are implemented in Italy.

Scenarios	Peak size (95% CI)	Peak time (95% CI)	Final size (95% CI)
$\beta$	5492 (5277, 5708)	Mar 26 (24, 27)	$2.59 \times 10^5$ ( $2.10 \times 10^5, 3.08 \times 10^5$ )
$0.8\beta$	4340 (4235, 4564)	Mar 25 (24, 26)	$2.03 \times 10^5$ ( $1.75 \times 10^5, 2.31 \times 10^5$ )
$0.6\beta$	3323 (3197, 3450)	Mar 25 (24, 26)	$1.52 \times 10^5$ ( $1.37 \times 10^5, 1.66 \times 10^5$ )
$q = 0$	5492 (5277, 5708)	Mar 26 (24, 27)	$2.59 \times 10^5$ ( $2.10 \times 10^5, 3.08 \times 10^5$ )
$q = 1/8$	3413 (3279, 3547)	Mar 20 (19, 21)	$1.22 \times 10^5$ ( $1.14 \times 10^5, 1.30 \times 10^5$ )
$q = 1/4$	2609 (2434, 2783)	Mar 18 (17, 19)	$8.51 \times 10^4$ ( $7.91 \times 10^4, 9.11 \times 10^4$ )

**Table 8**

The peak size, peak time and final size for Italy when varying node degrees of symptomatically infected individuals.

Degree	Peak size (95% CI)	Peak time (95% CI)	Final size (95% CI)
$k$	5492 (5277, 5708)	Mar 26 (24, 27)	$2.59 \times 10^5$ ( $2.10 \times 10^5, 3.08 \times 10^5$ )
$k - 1$	4583 (4410, 4755)	Mar 25 (24, 27)	$2.13 \times 10^5$ ( $1.81 \times 10^5, 2.44 \times 10^5$ )
$k - 2$	3683 (3546, 3821)	Mar 25 (24, 26)	$1.69 \times 10^5$ ( $1.50 \times 10^5, 1.88 \times 10^5$ )
$k - 3$	2795 (2680, 2910)	Mar 24 (23, 26)	$1.28 \times 10^5$ ( $1.17 \times 10^5, 1.39 \times 10^5$ )

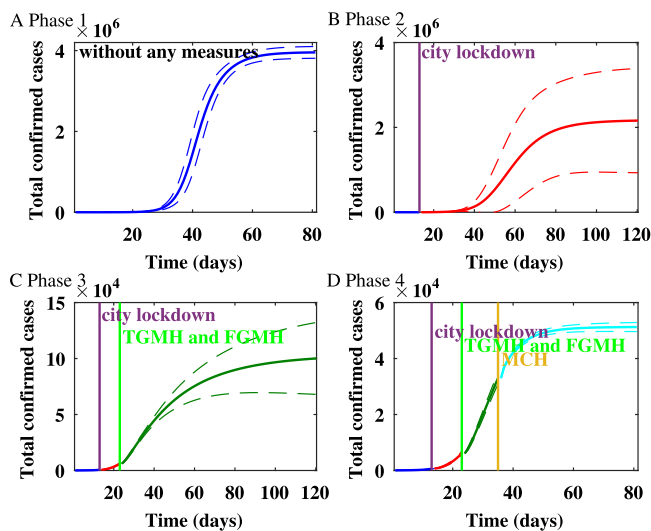


Fig. 4. The impact of the variability on the healthcare capacity on the spread of the epidemic in Wuhan on the Watts–Strogatz network.

show that the peak would have arrived earlier if the containment had been intensified.

When the probability of contact transmission coefficient  $\beta$ , is reduced by 20% by personal protection or social distancing, the peak occurs one day earlier, and the final epidemic size is reduced by 21.56%. When  $\beta$  is reduced by 41.44%, the peak occurs one day earlier, and the final epidemic size is reduced by around 40%.

When  $q = 1/8$ , the peak occurs six days earlier, and the final epidemic size is reduced by 52.87%. Yet, when  $q = 1/4$ , the peak occurs eight days earlier, and the final epidemic size is reduced by 67.12%. When the node degree of symptomatically infected individuals is reduced by 1, 2, and 3, the number of new cases produced per day at the peak is reduced by 16.50%, 32.93%, and 49.11%, respectively. The final epidemic size is reduced by 17.90%, 34.70%, and 50.51% when the node degrees of symptomatically infected individuals are reduced by 1, 2, and 3, respectively.

### 5. Summary and discussions

Modeling the dynamics of COVID-19 epidemics and assessment of mitigation strategies could be instrumental to public health agencies for surveillance and healthcare planning. For the models to be reliable, the simulated epidemic must account for the stochastic and heterogeneous contact among individuals. Hence, we developed a network model that captured the contact heterogeneity among individuals. We applied the

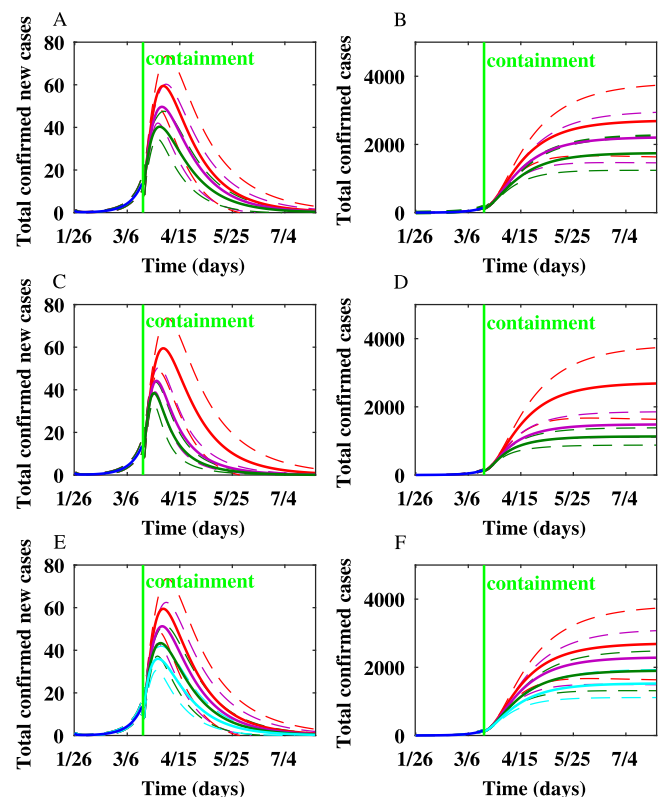
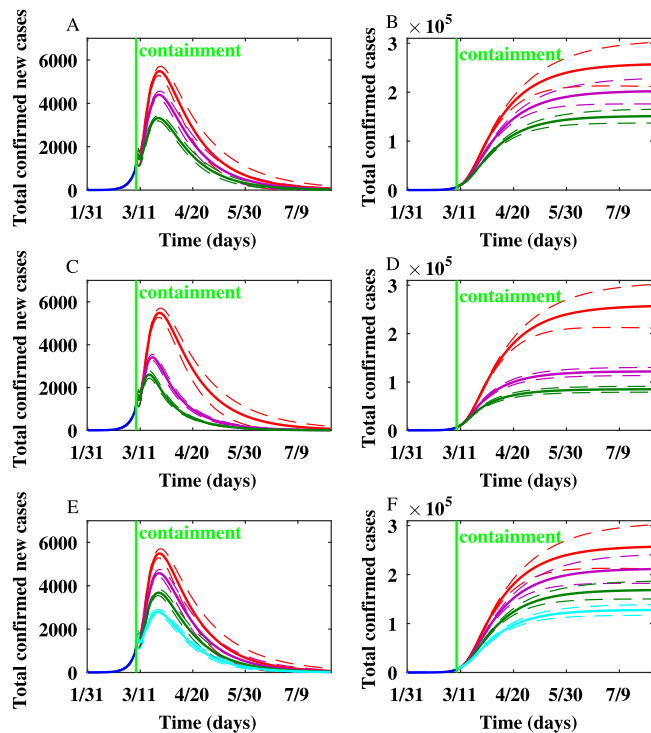


Fig. 5. The impact of mitigation strategies on the spread of COVID-19 epidemic in Toronto on the Watts–Strogatz network. In this figure and the following figure, the dashed lines represent 95% confidence intervals. In (A) and (B), the red, purple, and green lines represent that the transmission rates are unchanged, reduced by 20%, and reduced by 40%, respectively. In (C) and (D), the red, purple, and green lines represent the rate of quarantine,  $q = 0, 1/8, \text{ and } 1/4$ , respectively. In (E) and (F), the red, purple, and green, and light blue lines represent that the node degrees of symptomatically infected individuals are reduced by 0, 1, 2, and 3, respectively. (A) The number of newly infected individuals after reducing the transmission rates by personal protection and social distancing. (B) The cumulative number of infected individuals after reducing the transmission rates by personal protection and social distancing. (C) The number of newly infected individuals after close contacts are quarantined. (D) The cumulative number of infected individuals after close contacts are quarantined. (E) The number of newly infected individuals after the node degrees of symptomatically infected individuals are reduced. (F) The cumulative number of infected individuals after the node degrees of symptomatically infected individuals are reduced.

model to analyze the transmission potential, and mitigation strategies for curbing the spread of COVID-19 epidemics in the cities of Wuhan, China and Toronto, Canada, and in the Italian Republic. The epidemic threshold derived from our network model can be used to predict the



**Fig. 6.** The impact of mitigation strategies on the spread of COVID-19 epidemic in Italy on the Watts–Strogatz network. (A) The number of newly infected individuals after reducing the transmission rates by personal protection and social distancing. (B) The cumulative number of infected individuals after reducing the transmission rates by personal protection and social distancing. (C) The number of newly infected individuals after close contacts are quarantined. (D) The cumulative number of infected individuals after close contacts are quarantined. (E) The number of newly infected individuals after the node degrees of symptomatically infected individuals are reduced. (F) The cumulative number of infected individuals after the node degrees of symptomatically infected individuals are reduced.

risks of spreading scenarios. We also provided an explicit expression of the final epidemic size, which facilitates estimating the scale of an outbreak for any region of interest. Our results provide insights in defining a mathematical framework for the analysis and containment of epidemic transmission in the real world.

A wide range of mitigation strategies can be examined by the flexible model framework. It can be extended to quantify the effectiveness of personal protection, social distancing, reducing the node degree of infected individuals, and quarantine on the dynamics of epidemics in different regions. When the mitigation strategy is intensified, the model predicts that the number of new cases peaks earlier and the final epidemic size is greatly reduced.

The social contact network structure and parameter values determine the transmission and epidemic course of such an emerging infectious disease. We choose the Watts–Strogatz to approximate real social networks, when the exact contact tracing data is unavailable. We assumed that the range of the node degree is between one and ten for each network in the absence of real contact tracing data, that is, on average each day an infected person would have between one and ten contacts where they could transmit the infection to another person. In the real world, the range of the degree will depend on the distribution of the household sizes of the region and time being studied. Moreover, the network structure can be altered by behavior change of individuals during epidemics. When this happens, the network structure can be adapted in our model to predict the impact of these changes on the epidemic threshold, epidemic peak value, peak time, stopping time, and final size of infected population.

The epidemics for the three places under study were fitted very well by our model with a small confidence interval. Hence, the forecasts by the model can be reliable. We did not provide the stopping time since too many uncertainties may affect the duration of the epidemics. As shown in the simulations, the transmission dynamics for four phases in Wuhan are quite different due to the variability on the intensity of interventions, the availability of healthcare facilities, as well as the utilization of personal protective equipment (PPE). The dynamics in the first phase is quite different from that in the second phase for Toronto. The same phenomenon is observed in Italy.

At the early stage, almost no interventions were implemented, and the public was not aware of or did not pay much attention to the severity of the highly contagious disease. With the increase of the number of reported confirmed cases and with the aid of social media, the public becomes aware of the severe consequence and has increased the level of personal protection and have avoided gathering, so that the reproduction number decreases and the estimated epidemic size declines by reducing the node degree of the network. Similarly, after applying the mitigation measures in Italy on March 8 and closing all schools in Toronto on March 18, the epidemics tend to be under control.

Hence, social distancing, self-isolation, quarantine, the utilization of PPE, and other measures of avoiding exposure to the virus can greatly reduce the size of infection during the COVID-19 outbreak. Therefore, it is essential to raise the awareness of these countermeasures to avoid contact between individuals. The possibility of recurrent outbreaks of the disease cannot be overstated. Even if the number of new cases is declining, it is still necessary to continue taking protective measures to prevent the occurrence of future outbreaks. The social media should warn the public not to relax their vigilance against the contagion of such a highly infectious disease.

#### 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.

#### Acknowledgments

LX is funded by Fundamental Research Funds for the Central Universities of China. JCM received startup funding from La Trobe University. WS is funded by the National Science Foundation for Young Scholars of Heilongjiang Province, China QC2018004, and Fundamental Research Funds for the Central Universities of China. JGEF is supported by multidisciplinary, Mexico grant SIP-IPN 20196759. HZ is supported by Canadian Institutes of Health Research (CIHR), Canadian COVID-19 Math Modelling Task Force, and York Research Chair program of York University, Canada.

#### References

- [1] R. Anderson, R. May, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, Oxford, 1991.
- [2] A.L. Ziff, R.M. Ziff, Fractal kinetics of COVID-19 pandemic, 2020, medRxiv, <https://doi.org/10.1101/2020.02.16.20023820>.
- [3] H. Wang, Z. Wang, Y. Dong, et al., Phase-adjusted estimation of the number of coronavirus disease 2019 cases in Wuhan, China, *Cell Discov.* 6 (1) (2020) 1–8.
- [4] A.J. Kucharski, T.W. Russell, C. Diamond, et al., Early dynamics of transmission and control of COVID-19: a mathematical modelling study, *Lancet Infect. Dis.* 20 (5) (2020) 553–558.
- [5] M.U. Kraemer, C.H. Yang, B. Gutierrez, et al., The effect of human mobility and control measures on the COVID-19 epidemic in China, *Science* 368 (6490) (2020) 493–497.
- [6] M. Chinazzi, J.T. Davis, M. Ajelli, et al., The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak, *Science* 368 (6489) (2020) 395–400.
- [7] N. Ferguson, D. Laydon, G. Nedjati Gilani, et al., Report 9: Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand, 2020, <http://dx.doi.org/10.25561/77482>.

- [8] J. Hellewell, S. Abbott, A. Gimma, et al., Feasibility of controlling covid-19 outbreaks by isolation of cases and contacts, *Lancet Glob. Health* 8 (4) (2020) e488–e496.
- [9] S. Zhang, M. Diao, W. Yu, L. Pei, Z. Lin, D. Chen, Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the diamond princess cruise ship: A data-driven analysis, *Int. J. Infect. Dis.* 93 (2020) 201–204.
- [10] B.F. Maier, D. Brockmann, Effective containment explains subexponential growth in recent confirmed COVID-19 cases in China, *Science* 368 (6492) (2020) 742–746.
- [11] D. Welch, S. Bansal, D.R. Hunter, Statistical inference to advance network models in epidemiology, *Epidemics* 3 (1) (2011) 38–45.
- [12] C. Moore, M.E. Newman, Epidemics and percolation in small-world networks, *Phys. Rev. E* 61 (5) (2000) 5678.
- [13] S. Bansal, B.T. Grenfell, L.A. Meyers, When individual behaviour matters: homogeneous and network models in epidemiology, *J. R. Soc. Interface* 4 (16) (2007) 879–891.
- [14] M.E. Hochberg, Non-linear transmission rates and the dynamics of infectious disease, *J. Theoret. Biol.* 153 (3) (1991) 301–321.
- [15] C.L. Addy, I.M. Longini Jr, M. Haber, A generalized stochastic model for the analysis of infectious disease final size data, *Biometrics* 47 (3) (1991) 961–974.
- [16] D.J. Watts, S.H. Strogatz, Collective dynamics of small-world networks, *Nature* 393 (6684) (1998) 440.
- [17] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: Structure and dynamics, *Phys. Rep.* 424 (4–5) (2006) 175–308.
- [18] M. Boguna, R. Pastor-Satorras, A. Vespignani, Epidemic spreading in complex networks with degree correlations, in: *Statistical Mechanics of Complex Networks*, in: *Lecture Notes in Physics*, vol. 625, 2003, pp. 127–147.
- [19] R. Barrat, M.B.A. Vespignani, *Dynamical Processes on Complex Networks*, Cambridge University Press, Cambridge, 2008.
- [20] O. Diekmann, J.A.P. Heesterbeek, J.A. Metz, On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations, *J. Math. Biol.* 28 (4) (1990) 365–382.
- [21] P. Van den Driessche, J. Watmough, Further notes on the basic reproduction number, in: *Mathematical Epidemiology*, Springer, 2008, pp. 159–178.
- [22] Y. Wang, J. Cao, A. Alofi, A.M. Abdullah, A. Elaiw, Revisiting node-based sir models in complex networks with degree correlations, *Physica A* 437 (2015) 75–88.
- [23] Y. Wang, J. Cao, G. Huang, Further dynamic analysis for a network sexually transmitted disease model with birth and death, *Appl. Math. Comput.* 363 (2019) 124635.
- [24] H. Haario, M. Laine, A. Mira, E. Saksman, DRAM: efficient adaptive MCMC, *Stat. Comput.* 16 (4) (2006) 339–354.
- [25] Health Commission of Hubei Province, 2020, <http://wjw.hubei.gov.cn/bmdt/ztl/fkxgzbdgrfyyq/>. (Accessed 7 April 2020).
- [26] World Health Organization, 2020, <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/>. (Accessed 7 April 2020).
- [27] Hubei Provincial Bureau of Statistics, 2020, Available from: <http://tj.hubei.gov.cn/tjsj/sjksx/tjn/qstjn/>. (Accessed 17 February 2020).
- [28] F. Zhou, T. Yu, R. Du, et al., Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study, *Lancet* 395 (10229) (2020) 1054–1062.
- [29] J. Qiu, Covert coronavirus infections could be seeding new outbreaks, *Nature* (2020) <https://www.nature.com/articles/d41586-020-00822-x>.
- [30] Toronto Population 2020, 2020, <https://worldpopulationreview.com/world-cities/toronto-population/>. (Accessed 17 March 2020).
- [31] World Health Organization, 2020, <https://www.who.int/countries/ita/en/>. (Accessed 27 March 2020).