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.
Spread of COVID-19 in China: analysis from a city-based epidemic and mobility model

Ye Wei a, Jiaoe Wang b,c,*, Wei Song d, Chunliang Xiu e, Li Ma c,f, Tao Pei c,f

a Key Laboratory of Geographical Processes and Ecological Security in Changbai Mountains, Ministry of Education, School of Geographical Sciences, Northeast Normal University, Changchun, Jilin 130024, China
b Key Laboratory of Regional Sustainable Development Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
c College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China
d Department of Geography and Geosciences, University of Louisville, Louisville, KY, USA
e State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
f College of Jang Ho Architecture, Northeastern University, Shenyang, Liaoning, China

ARTICLE INFO

Keywords:
City lockdown
COVID-19
Epidemic model
Containment measure
Population mobility

Abstract

Understanding the processes and mechanisms of the spatial spread of epidemics is essential for making reasonable judgments on the development trends of epidemics and for adopting effective containment measures. Using multi-agent network technology and big data on population migration, this paper constructed a city-based epidemic and mobility model (CEMM) to simulate the spatiotemporal of COVID-19. Compared with traditional models, this model is characterized by an urban network perspective and emphasizes the important role of intercity population mobility and high-speed transportation networks. The results show that the model could simulate the inter-city spread of COVID-19 at the early stage in China with high precision. Through scenario simulation, the paper quantitatively evaluated the effect of control measures “city lockdown” and “decreasing population mobility” on containing the spatial spread of the COVID-19 epidemic. According to the simulation, the total number of infectious cases in China would have climbed to 138,824 on February 2020, or 4.46 times the real number, if neither of the measures had been implemented. Overall, the containment effect of the lockdown of cities in Hubei was greater than that of decreasing intercity population mobility, and the effect of city lockdowns was more sensitive to timing relative to decreasing population mobility.

1. Introduction

Epidemics refer to infectious diseases that can infect a large number of people, such as influenza, meningitis, cholera, and COVID-19. Serious epidemics can cause the deaths of tens of thousands of people in a short time (Anderson & May, 1992). Since Hamer and Ross’s pioneering work that quantitatively investigated the spread of measles, scientists have been committed to developing mathematical models for a better understanding of the epidemic transmission process (Anderson & May, 1992). The SIR compartment model and its variants (such as SEIR and SIRS) are the most commonly used models for analyzing the propagation dynamics of epidemics (Kermack & McKendrick, 1927; Murray, 1993; Small & Tse, 2005). These models are widely used to estimate the effective reproduction number and forecast the peak duration or the inflection point of epidemics and constitute the mainstream of epidemic dynamics research (Gu et al., 2020; Lai et al., 2015; Lai et al., 2020; Paine et al., 2010; Read et al., 2020; Wu et al., 2020). These models, however, have tended to minimize the impact of geographic heterogeneity, and they share a common weakness in that related spatial aspects of the spread of epidemics are largely ignored (Arino et al., 2007). For instance, when running an SIR simulation for the whole country, it is difficult to know an epidemic’s severity in each city. It would still be unreasonable even if a specific SIR model were designed for each city. There is at least a system openness problem here. Since each city is an...
open system, the impact of external input cannot be ignored. Especially with advancements in transportation technologies and a well-developed transport network, population mobility between cities is becoming more and more convenient (Zhang et al., 2009; Zhang et al., 2020).

In order to solve this problem, researchers have proposed a meta-population model to characterize the long-distance movement of individuals between cities and distinguish them from intracity travel behavior (Colizza et al., 2007). In this model, based on the geographic location, the whole population is divided into several subgroups linked by population movement to form a metapopulation network. Individuals are likened to particles, and particles move randomly on the metapopulation network, including the movement within and between sub-populations (Colizza et al., 2007). Subsequent research revealed that the movement of individuals is not random, but purposeful (Tang et al., 2009). In other words, the accuracy of the metapopulation model will be greatly increased if ancillary population flow data are available and used to define the purpose and probability of each individual flow. Furthermore, the metapopulation model is also built on the simulation of individual behavior. For this reason, when the size of a population is large or detailed population mobility data are included, the model could become complex and require large computing resources. The real-time calibrations could even require supercomputing resources (Tizzoni et al., 2012), which makes application of this kind of model difficult. In addition, population mobility is often simulated with a gravity model and radiation model rather than the actual measured population flow data. This could affect the accuracy of the modeling to some extent.

According to their principle and mechanism, the compartment models can be seen as macro models, and the metapopulation models can be regarded as micro models. Each has its own advantages and disadvantages in answering the different questions mentioned above. Therefore, it is supposed that a compromise scale might make full use of the advantages and avoid the disadvantages. In this respect, some scholars have advocated using spatial networked models to simulate the spread of epidemics by incorporating human mobility patterns generated with big data from mobile phones and global flight networks (Brockmann & Helbing, 2013; Chinazzi et al., 2020; Lai et al., 2020).

The biggest benefit of using spatial networked models is that they derive the temporal and spatial information of the epidemic simultaneously, but their value is restricted in practice by the large amount of calculations and lack of reliable, spatially detailed, timely population mobility data.

In this study, we built a City-based Epidemic and Mobility Model (CEMM) using multi-agent network techniques to simulate the spatial spread of epidemics. Different from existing models, the model emphasizes a city network perspective and the important role of intercity population mobility and fast transportation networks for epidemic transmission. It would be helpful to understand the spatiotemporal mechanism of epidemic transmission and formulate corresponding policies from a geographic network perspective for containing the epidemic. This model was applied to rebuilding the spatiotemporal spread process of COVID-19 among cities in China during the early stage of the outbreak and evaluating the effect of containment measures under various scenarios.

The remainder of this paper is organized as follows. Section 2 introduces the theoretical basis of the model. Section 3 discusses the proposed model in detail. Section 4 introduces the research case and data sources. Section 5 presents modeling results and accuracy evaluations against empirical epidemic data. Finally, discussions and concluding remarks are provided in Section 6.

2. Theoretical basis of the model

2.1. Principle of spatial transmission of epidemics

According to previous studies (Brockmann & Helbing, 2013; Gao et al., 2020; Wang & Li, 2014), epidemics not only prevail intracity, each of which can be regarded as a specific population, but are also transmitted from city to city through the travel of infected individuals. For epidemic transmission between cities, most earlier models or studies focused on the output of the epicenter city (Brockmann & Helbing, 2013; Wu et al., 2020; Zhao et al., 2020), while few discussed the output of other cities. This type of transmission could thus be considered a monocentric spreading mode (Fig. 1(a)).

With well-developed high-speed transport networks in many countries, especially in China (Jiao et al., 2014; Jiao et al., 2017), people can travel from one city to another within a short time. By the end of 2019, the total mileage of China’s high-speed railway reached 35,000 km, ranking first in the world. In 2019, 660 million flights and 2.3 billion high-speed train rides were taken, accounting for 16.8% of all transportation in China. Meanwhile, the Chinese central government announced the plan “Building China into a Country with a Strong Transportation Network”, which requires major cities to be reachable within 3 h by 2035. Therefore, traveling between cities by high-speed train services or air flights in China is quite easy and population mobility is more improved than ever before. As such, the input of infectious cases in a certain area might not have originated from the epicenter, but could have come from other areas following a polycentric or networked spreading mode (Fig. 1(b)).

With the intensification of an epidemic, the second or the third generation of infections could increase, making the polycentric spreading mode more apparent and more reasonable. For example, among the six imported COVID-19 cases reported in Beijing on 10 March 2020, five cases were from Italy and one case was from the USA. Kraemer et al. (2020) found that 43% of the early cases (before 23 January 2020) reported outside of Wuhan did not have a known travel history to Wuhan and these were distributed across China. Accordingly, this research will employ the polycentric spreading mode as the basic hypothesis to build the model and perform analysis.

2.2. Population mobility and inter-city epidemic spread

Many existing studies explore the relationship between population mobility and epidemic transmission. As early as 1963, Prosorè studied the relationship between population mobility and trypanosomiasis prevalence in Africa (Prothero, 1963). Colizza et al. (2007) and Khan et al. (2009) demonstrated that international air travel provided good prediction for the worldwide spread of SARS and influenza A/H1N1 2009. Further, Agaard-Hansen et al. (2010) revealed that population movement is a key factor for explaining the spread of tropical epidemics. Subsequently, some empirical studies have confirmed that population mobility was closely related to the prevalence and spread of epidemics (Schlagenhauf et al., 2014; Zhang et al., 2015). In principle, if an epidemic exists in a city, and if the population travels outside the city, then there would be a certain amount of people infected, as per a certain ratio. The larger the population flow, the more infected cases there will be. In other words, the magnitude of population mobility between cities represents the probability of epidemic transmission (Brockmann & Helbing, 2013). Therefore, population mobility is regarded as an important factor for influencing intercity transmission of the epidemic in the model built in this paper. Since the outbreak of COVID-19, scholars have explained the spatial spread of the epidemic from the perspective of population mobility. Kraemer et al. (2020) used real-time mobility data from Wuhan and detailed cases to ascertain the impact of control measures in China and confirmed that the spatial distribution of COVID-19 cases was explained well by human mobility in the early spread of the epidemic. Jia et al. (2020) argued that population migration out of Wuhan before January 23, 2020, greatly influenced the spatiotemporal distribution of confirmed COVID-19 cases in China. However, Shi and Liu (2020) argued that the internal migrants during the Chinese Spring Festival in China should be held less responsible for the initial spread of COVID-19 than business and tourism connections with Wuhan. Although they all agree that population mobility played an important
role in the spread of COVID-19 in China, existing studies mainly used the population flows out of Wuhan to simulate the spread of COVID-19. The reality, however, is that people in China can travel quickly from Wuhan to another city, and then further to a third city in a short time due to the well-developed high-speed transport network. As such, we need to calculate human mobility from a complex and interactive network for simulating the spread of COVID-19.

3. Methodology

3.1. Basic model: CEMM

As mentioned in Section 2.1, the increase in infectious cases in each city resulted from two primary processes: transmission from other cities and contagion inside the city (Fig. 1(b)). Thus, our epidemic modeling involves these as its components (equations).

3.1.1. Intercity transmission

In this research, cities are the basic unit of analysis and nodes of the population mobility network. Each node has basic attributes such as population size and number of infectious cases during the epidemic. As the spread of the epidemic among nodes is mainly affected by the infectivity of the virus, population mobility scale, and the number of infectious cases, the intercity epidemic transmission component is formulated as shown in Eq. (1):

$$N_{ij,t}^+ = \frac{N_{ij,t-1} \times \beta \times P_{ij}}{\text{popu}_i} \quad (1)$$

where $N_{ij,t}$ is the number of infectious cases in city $i$ at time $t$; $N_{ij,t-1}$ accounts for the initial seed, that is, the number of infectious cases in city $i$ at the time $t$; $P_{ij}$ is the number of people traveling from city $i$ to city $j$ every day; $\text{popu}$ is the total population in city $i$; and $\beta$ is the epidemic rate (i.e., the number of infectious cases per time point). The epidemic rate can be inverted according to the growth of the number of infectious cases in the epicenter. Specifically, we constantly ran the model and adjusted the $\delta$ value until the model fit the growth in the number of infectious cases inside and outside the epicenter both approached the numbers of actual infectious cases.

3.1.2. Intracity contagion

In addition to transmission among cities, the epidemic also spreads between individuals within a city. The intracity contagion component is shown in Eq. (2):

$$N_{ij,t}^- = N_{ij,t-1} \times \beta \times \delta \quad (2)$$

where $N_{ij,t}$ represents the increase in infectious cases via internal diffusion in city $j$ at time $t$; $N_{ij,t-1}$ is the number of infectious cases at time $t$; $\beta$ is the epidemic rate from Eq. (1); and $\delta$ is the reduction factor ranging from 0 to 1.

For the epicenter, the value of $\delta$ is 1 or less than 1. This reflects the proposition that other cities would have lower severity of the epidemic than the epicenter, as they could follow proven and tested control measures for containing epidemic diseases, as latecomers. The $\delta$ value of other cities outside the epicenter could be determined by repeat tests. Specifically, we constantly ran the model and adjusted the $\delta$ value until the total numbers of infectious cases inside and outside the epicenter both approached the numbers of actual infectious cases.

3.1.3. Estimating infectious cases in each city

Combining the intercity and intracity components, the estimated number of infectious cases in city $j$ at time $t$, i.e., $N_{ij,t}$ can be calculated using Eq. (3):

$$N_{ij,t} = N_{ij,t-1} + \sum_i N_{ij,t}^- + N_{ij,t}^+ \quad (3)$$

Using Eqs. (1) through (3), the model can predict the number of infectious cases in each city on a specific day. We refer to this model as the “City-based Epidemic and Mobility Model” (CEMM) in this paper.

3.2. Improved model: adapted CEMM

A preliminary simulation analysis using CEMM revealed that the number of infectious cases in cities with larger population inflows tended to be overestimated, while the number of cases in cities with lesser inflows tended to be underestimated. Furthermore, the number of infectious cases in some cities distant from the epicenter tended to be underestimated, and the number of cases in nearby cities tended to be underestimated. To address these issues, the basic CEMM model was improved by considering the two effects.

3.2.1. Effect of decreasing returns to mobility scale

The power function is often used to measure the increasing or decreasing law of returns to scale in economics (Szakolczai & Stahl, 1969). The famous Cobb-Douglas Production Function is a typical application. The power function was used to simulate the effect of decreasing returns to mobility scale in this study. The component of intercity transmission originally formulated in Eq. (1) was revised to the following formula:

$$N_{ij,t}^+ = N_{ij,t-1} \times \beta \times \left(\frac{P_{ij}}{\text{popu}_j}\right)^{\gamma} \quad (4)$$

where $\gamma$ is the influence coefficient of the flow scale. For each city, the input of the city-based Epidemic and Mobility Model (CEMM) in this paper.
of the migrating population on the epidemic situation increases at a lower rate when compared to the increase in the migration scale. This could explain the reason why the number of infectious cases of COVID-19 in some cities with a large population inflow is not as high as expected.

3.2.2. Effect of distance decay

In order to simulate the distance decay effect, a Gaussian function (Dai, 2011) was introduced, as shown in Eq. (5):

$$ G(d_{ij}, d_0) = \frac{1}{1 - e^{-\frac{d_{ij}}{d_0}}} $$

(5)

where $d_{ij}$ is the geometric distance between city i and city j, and $d_0$ is a constant, which is the geometric distance between the farthest two cities in China. From this we get Eq. (6).

$$ n_{ij}^\text{expected} = \frac{N_{t-1} \times \beta \times \frac{P_i \times G(d_{ij}, d_0)}{\text{pop}_i}}{1 - \epsilon} $$

(6)

By using Eq. (6) instead of Eq. (3) for the intercity transmission component, we get a revised and improved model, which we refer to as the “Adapted CEMM”.

3.3. Multi-agent network as a technical support

The multi-agent technique has been widely used in epidemic simulation (Eubank et al., 2004; Ferguson et al., 2005; Merler & Ajelli, 2010), but in existing multi-agent models, the agents typically represent single individuals, and the epidemic spread between individuals is through contact with each other. Such models are effective in simulating the impact of human behavior and social network structure on the spread of an epidemic, but when the population reaches a certain size, the computational cost and the demand for detailed data can be a major problem (Balcan et al., 2010). With the city being the basic analytical unit, this study assumes each city as an agent node, and uses the total population, the number of infectious cases, and other parameters of the city as node attributes. The population mobility flows between cities are considered the connections between nodes to build a multi-agent network, also known as the city network. Thus, the spatial spread process of the epidemic can be simulated as the growth in the number of infectious cases in the city and the transmission of a virus among cities across the city network. The specific simulation process was implemented using Netlogo software, as described by Elizabeth (2007).

4. Case study and data sources

4.1. Spread of COVID-19 in China and its control measures

The novel coronavirus—severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Coronaviridae Study Group of the International Committee on Taxonomy of Viruses (2020))—has been identified as the cause of the 2019 coronavirus disease (COVID-19) respiratory illness that was first detected in Wuhan, China, in late 2019. By 8 March 2020, more than one hundred countries and regions had reported confirmed cases of COVID-19 and a total of 105,586 confirmed cases were reported globally as of 6:00 CET (0500 UTC), with 3584 deaths. On 11 March, the World Health Organization declared the novel coronavirus outbreak a pandemic.

Following the outbreak of COVID-19 in Wuhan, this highly contagious virus quickly spread, with new daily cases peaking at over 3887, and with more than 320 prefecture-level cities in China reporting infectious cases on 4 February 2020. Since the outbreak of the virus, the central and local governments of China have taken strict, effective measures to prevent the spread of coronavirus, including reporting epidemic notifications on time, active treatment and isolation of patients, encouraging people to wear face masks when going out, instructing people to practice social distancing, wash hands often, and cover their mouth and nose while coughing and sneezing. At the same time, some unprecedented interventions have been implemented—including a city lockdown in Hubei Province, travel restrictions, and blocking transport connections between cities, towns, communities, and villages (Gu et al., 2020). The government encouraged everyone to stay at home and reduce unnecessary travel.

With the global spread of the COVID-19 epidemic, various countries gradually began to adopt a similar strategy, including Italy, Spain, and other countries, which included the closure of national or urban borders. In China, according to the Ministry of Transport, the passengers carried by railways, highways, and airlines across the country dropped by more than 76.4% during the lockdown period as compared to last year. These control measures may have an impact on the model’s predictions and were accordingly considered in the model.

4.2. Data source

Tencent migration data was used as a measurement of population mobility between cities. The data was obtained based on LBS (Location Based Services) technology that records the daily population flows between cities in a full, dynamic, real-time, and intuitive manner (Pan & Lai, 2019). Before current data on population mobility is available, data from previous years during the same period is a valuable resource that can be used as a suitable substitute. Therefore, the Tencent migration data from 2019 was used as an approximate substitute. In order to keep pace with the epidemic based on the lunar calendar, which has implications for holidays and travel, the daily mobility data of 362 cities in China from 21 January (the sixteenth day of the twelfth lunar month of the previous year) to 9 February (the fifth day of the first lunar month) in 2019 were derived from the Tencent Location big data platform (https://heat.qq.com/index.php). Using this data, we constructed a directed inter-city population mobility network (Fig. 2), as per the method given by Wei et al. (2018).

The total population of each city at the end of 2018 was obtained from the China city statistical yearbooks and the statistical yearbook of each province. The number of confirmed COVID-19 cases from 17 January 2020 to 6 February 2020 was collected from the official data released by the National Health Commission of China.

5. Accuracy evaluation and result analysis

5.1. Accuracy evaluation

According to the officially announced data, the number of newly confirmed cases of COVID-19 outside Hubei Province reached a peak on 4 February 2020 and then decreased constantly from that point on due to the comprehensive containment measures. In order to evaluate the accuracy of the model, the number of COVID-19 infectious cases in each city from 26 January 2020 to 4 February 2020 were calculated separately using both CEMM and Adapted CEMM.

Notably, there is a time delay between the onset of COVID-19 symptoms and the reporting of cases, which was revealed by previous studies (Lai et al., 2020; Xu et al., 2020). Therefore, the number of reported confirmed cases cannot be used directly to evaluate the accuracy. Lai et al. (2020) summarized the average days from illness onset to the report of the first case by each county in China. They found that the average interval from the onset of COVID-19 to confirmation dropped from 12 days in the early stages of the outbreak to 2 days after 28 January 2020, and presented this information as a supplementary table. According to the table and daily reported cases, we approximately calculated the total number of infections per day for each city in China. For example, if the number of reported confirmed cases on 6 February in...
city x was 100, then the number of infectious cases on 4 February in city x was also deemed to be 100, since the time delay of onset and confirmation was 2 days.

Subsequently, linear regression analysis was carried out between the simulated value of the number of infectious cases and the officially announced real value in each city. The scatter plots are shown in Fig. 3; adjusted R-squares, standard error, and significances are shown in Table 1. Since cities in Hubei Province were in a lockdown, the number of infectious cases was much larger than that of cities outside; error analysis was carried out separately for all cities in the country, cities in Hubei Province, and cities outside the province. In addition, data from Wuhan was eliminated to avoid the overfitting issue in the regression analysis and scatter plots.

From Table 1, we can see that the overall adapted CEMM has higher R-squares and lower standard errors than CEMM, especially for cities outside Hubei Province, and the highest fitting accuracy (R-square) reached 0.854 for all cities in China. Fig. 3 shows the variance of the scatter plots and R-squares from 26 January to 4 February 2020. The fitting accuracy (R-square) of the model was found to increase continuously. The positions of trend lines of the scatter plots were adjusted continuously, and gradually tended to be stable after 30 January. In terms of the obvious changes in trend lines before 30 January, the possible reason is that the number of cases in many cities was too small and thus it was difficult to discern the real law of spatial distribution of the epidemic in the earlier stage due to the insufficient sample size. With the spread of the epidemic, the number of confirmed cases increased and the sample size gradually became sufficient, which in turn led to a better fitting effect on the model. Overall, the model has adequate precision, accuracy, and capability for describing the spatial diffusion of the epidemic, especially after the epidemic reached a certain scale.

5.2. Analysis of simulation results

5.2.1. Geo-visual result analysis

Visualization analysis was carried out using ArcGIS software. Simulated and actual distribution of COVID-19 cases are shown in Figs. 4 and 5, respectively.

The spatial distribution of the COVID-19 epidemic does not simply present a central-periphery pattern as expected, whether from the real results (Fig. 4) or from the simulated results (Fig. 5). Due to the greater number of infectious cases in the three largest urban agglomerations in China—Beijing-Tianjin-Hebei, Yangtze River Delta, and Pearl River Delta—the central-periphery pattern broke up, and the characteristics of leap-frogging diffusion formed. Such a pattern shows that neither population mobility nor distance alone can explain the spread of the epidemic.

In terms of geographical distribution, there is a significant difference in the severity of the epidemic on both sides of Hu’s line (Heihe-Tengchong Line) (Hu, 1935), which is an important geographic dividing line in terms of population density as well as natural conditions in China. The number of infectious cases to the east of the line is much larger than that on the other side of the line, which is in step with the spatial distribution pattern of population and population mobility during the Chinese Spring Festival (Wei et al., 2018).

5.2.2. Performance of “lockdown of cities” and “decreasing population mobility” policies to contain the COVID-19 epidemic

Starting on 23 January 2020, people were not allowed to enter or leave Wuhan in order to contain the rapid spread of the COVID-19 epidemic. An hour and 20 min later, Ezhou, another city in Hubei, announced that it had shut down the railway station and limited the entry and exit of people. The next day (24 January), eight prefecture level cities in Hubei (e.g., Huangshi) implemented similar measures. On
Fig. 3. A comparison of estimated cases vs. actual infectious cases in cites before 4 February 2020.
28 January, Xiangyang cut off external traffic. Furthermore, residents were not allowed to travel between cities, counties, and communities in Hubei. Media referred to this measure as a “lockdown of cities”. At the same time, due to traffic control, media publicity, a holiday extension, and other factors, the scale of the population flow between cities in other regions was also greatly reduced. The effects of these two factors can be simulated by.

decreasing the magnitude of intercity population migration. Specifically, population movements \( (P_{ij}) \) between all cities were reduced by 76.4% to simulate the decreased population mobility. The population flows of cities in Hubei Province were decreased to 0.0001 to simulate the city lockdown measures implemented there.

In order to evaluate the efficacy of the application and timing of the “city lockdown” policy in Hubei and the “decreased population mobility” policy on the transmission of the COVID-19 epidemic, ten scenarios were simulated, in which scenario 0 was the actual situation. The details of the scenario settings are as follows.

Based on the above ten scenarios, the growth in the number of infectious cases was simulated for all cities in the country, cities in Hubei Province, and cities outside Hubei Province (Fig. 6 and Table 2).

From Fig. 6 and Table 3, it can be seen that the city lockdown and decreased population mobility policies effectively contained the spread of COVID-19 in cities in Hubei Province and also in cities outside Hubei Province. The combination efficacy of city closure and migration flow reduction was far greater than that of a single measure. In this regard, it is evident that the decisive measures taken by China, such as shutting down cities in Hubei and cutting population mobility, were necessary to control the epidemic.

By applying the two measures, the total number of infectious cases in China as of 4 February 2020 was theoretically reduced by 77.56% (107,665 cases) compared to a scenario where neither of the policies were implemented. If both the measures had been implemented one week earlier, the total number of infectious cases in China could have been reduced by 88.97% (123,511 cases). However, if both the measures had been delayed one week later, the total number might only have been reduced by 47.06% (65,332 cases). The comparison between scenario 8 and scenario 9 suggests that the effect of city lockdowns in Hubei on the spread of the epidemic was stronger than that of decreased population mobility.

The curves of scenario 2 and scenario 6 are almost coincident, which means that the control effect was not improved much by restricting intercity population flow one week earlier than the actual time, under the premise of a city lockdown one week earlier. We also found that the difference in the number of cases between scenario 2 and scenario 3 was larger than that in scenario 4 and scenario 5. These findings indicate that the growth of infections was more sensitive to the timing of city lockdowns in Hubei than to the decreased population mobility.

### 5.2.3. Comparison of the roles of intercity transmission and intracity contagion

According to the principle of CEMM (Fig. 1(b)), the increase in the number of infectious cases is sourced from two parts—contact infection within each city and input from other cities. With our model, one can roughly estimate how many infectious cases in each city come from internal contact infection, and how many from external input. This

| Analysis unit          | Coefficients | CEMM | Adapted CEMM |
|------------------------|--------------|------|--------------|
| All cities             | Adjusted R²  | 0.735-0.879 | 0.780-0.854 |
|                        | Std. error   | 0.020-0.057 | 0.016-0.035 |
|                        | Sig.         | 0.000    | 0.000        |
| Cities in Hubei Province| Adjusted R²  | 0.861-0.926 | 0.806-0.869 |
|                        | Std. error   | 0.057-0.183 | 0.051-0.111 |
|                        | Sig.         | 0.000    | 0.000        |
| Cities outside Hubei Province| Adjusted R² | 0.457-0.527 | 0.480-0.580 |
|                        | Std. error   | 0.066-0.113 | 0.036-0.059 |
|                        | Sig.         | 0.000    | 0.000        |

Table 1

R² and significance of correlation analysis between estimated and confirmed cases in cities.

---

Fig. 4. Spatial distribution of estimated cases by adapted CEMM in China on 4 February 2020.
function has rarely been reported in previous studies. As an experiment, we calculated the ratio of internal contact infectious cases and external input cases for each city. Due to the limited space, the data of each city could not be presented, but the analysis of the overall situation of the whole country is presented.

According to our model, as of 4 February, about 33.55% of the cases of infection in all cities (except Wuhan) came from intercity transmission, and 66.45% from intracity contagion. Fig. 7(a) shows the time-varying curve of the total number of cases of intercity transmission and intracity contagion from 10 January to 4 February 2020. Fig. 7(b) shows the obvious influence of city lockdowns in Hubei and decreased population mobility on the intercity transmission of the epidemic. After the implementation of the two measures on 22 January, the intercity transmission was retarded to a certain extent. In contrast, the growth trend of intracity contagion was not significantly affected.

### Table 2
Settings for ten scenarios.

| Scenarios | Containment measures |
|-----------|----------------------|
|           | Lockdown cities in Hubei | Decreased population mobility |
| Scenario 0 | Applied at actual time | Applied at actual time |
| Scenario 1 | Not applied | Not applied |
| Scenario 2 | Applied one week earlier | Applied at actual time |
| Scenario 3 | Applied one week later | Applied at actual time |
| Scenario 4 | Applied at actual time | Applied one week earlier |
| Scenario 5 | Applied at actual time | Applied one week later |
| Scenario 6 | Applied one week earlier | Applied one week earlier |
| Scenario 7 | Applied one week later | Applied at actual time |
| Scenario 8 | Not applied | Applied at actual time |
| Scenario 9 | Applied at actual time | Not applied |

### Table 3
Estimated number of infectious cases on 4 February 2020 under different scenarios.

| Cases in Hubei | Cases in other provinces | Cases in China |
|---------------|--------------------------|----------------|
| Scenario 0    | 19,141                   | 12,017         | 31,159         |
| Scenario 1    | 39,654                   | 99,169         | 138,824        |
| Scenario 2    | 12,731                   | 26,966         | 15,428         |
| Scenario 3    | 23,430                   | 20,323         | 43,754         |
| Scenario 4    | 15,144                   | 43,754         | 43,754         |
| Scenario 5    | 19,367                   | 21,677         | 41,045         |
| Scenario 6    | 12,731                   | 2581           | 15,313         |
| Scenario 7    | 29,476                   | 44,016         | 73,492         |
| Scenario 8    | 26,714                   | 29,230         | 55,945         |
| Scenario 9    | 19,367                   | 33,978         | 53,345         |
Comparing the changes of the two curves, before 26 January, the growth of infectious cases in each city mainly came from external input; after that date, the growth of infectious cases in each city came mainly from intracity contagion, and the proportion of infectious cases caused by intracity contagion kept rising. We also simulated the changing process of the curves under the scenario of no city lockdowns in Hubei and no decreased population mobility, as shown in Fig. 7(b). It turned out that if these two containment measures had not been taken, the situation would be exactly the opposite—intercity transmission would replace intracity contagion as the main cause of infectious case growth. This situation would not only accelerate the increase in the number of infectious cases, but also bring uncertainty and difficulty to the containment of the epidemic.

6. Discussion and conclusions

6.1. Discussion

Since randomness and uncertainty are the fundamental properties of nature, prediction is naturally difficult. Furthermore, the internal and external containment efforts and response times of each city are different, resulting in a reduction in the degree of intercity population flow, while the intracity diffusion index is also significantly different. In order to simplify the model, we did not consider these factors for each city separately. Even if we set containment parameters for each city separately, it is difficult to quantify these and we are faced with an a posteriori problem.

Moreover, according to general experience, under the premise of lack of effective control, the number of daily virus infections tends to increase exponentially. There may be a huge difference in the number of infectious cases before and after a two-day gap, and the difference in time will also increase the prediction error.

In addition, the diagnosis of infection is a complex process that includes biochemical and imaging evidence, because of which the number of reported cases cannot fully reflect the actual number of infections, which is also one of the important factors affecting the accuracy of the model. Some accidental factors are not included here, such as clustering infection and super-spreaders.

In summary, it is difficult to accurately predict the number of infectious cases in a city at a certain time due to many types of factors. However, the CEMM model has obvious value in explaining the spatiotemporal spread mechanism and in analyzing the development trend of epidemics.

From the principle of the model and the results of the simulated scenarios, we learned it is key to control the outputs of infectious cases from the epicenter, and it is also important to prevent the secondary transmission from other regions. Both “lockdown of cities” and “decreasing population mobility” policies had significant effects on containing the COVID-19 epidemic at its earlier stage. In other words, epidemic containment is a systematic “network project” that requires the joint efforts of all cities and regions to completely control the epidemic.

Due to limited space, we cannot discuss the specific mechanism of the effect of decreasing returns to scale and distance decay in detail. In future research, the complex network method will be used to deeply explore the impact of the social network structure and contact mode of the floating population on the infection capacity.

At the same time, research on the spatiotemporal pattern of population mobility and how to respond to crisis events will be carried out for different modes of transportation. The prediction model also requires a cure interface, and further consideration could be given to the influence of treatment factors on the development of epidemics in the future.

The results of the study were only for the spread process of an epidemic between cities in a country; the spread of the epidemic between transnational cities and countries was not part of the analysis. Due to the influence of entry-exit inspection, extra-long geographic distance and other factors, there may be different spatial laws when these factors are also considered, which in turn need further investigation.

6.2. Conclusions

In this study, a City-based Epidemic and Mobility Model (CEMM) was proposed by using multi-agent technology and big data on population migration. A case study of COVID-19 demonstrated that the model was useful in the simulation of the intercity transmission of an epidemic and could provide more geospatial information compared with traditional compartment models. The model suitably explains the spatial and temporal evolution of the COVID-19 epidemic in its initial stages. The main findings are as follows.

This study confirmed that population mobility can explain the mechanism of intercity transmission of the epidemic to a certain extent, and that the model can effectively predict the spread process of the epidemic with accurate population mobility data. However, the spread of the epidemic is not merely a central–radial structure spreading from the epicenter, but a complex network process. That is, cities should not only prevent the epidemic from the epicenter, but also protect against...
infection from other areas. Especially for countries with high-speed, large-volume transportation facilities and high mobility efficiency, the time distance between cities has been greatly shortened, and the population can travel between cities quickly and flexibly, in large numbers. This increases the complexity of the epidemic transmission, and the network effect should be paid more attention.

In this study, we found that population mobility has a decreasing return to the population mobility scale effect on epidemic transmission performance. It is generally believed that the floating population and the total population of the city should have the same infection rate and transmission efficiency, and the number of output infectious cases should have a linear relationship with the amount of the population migration. However, the results showed that the number of output infectious cases does not increase proportionally with the growth in population flow but shows a certain degree of decline. This phenomenon could be explained by individual networking, i.e., social networks. A social network is generally considered as a six-degree network, which is a typical Erdős–Rényi (ER) network (Erdős & Rényi, 1961). When all people are advised to be isolated at home and access to public places is restricted, this feature becomes more obvious. Assuming that the virus infects ER networks with a different number of nodes with the same probability of infection, it can be found that the virus spreads more slowly with the increase in the number of nodes, which could be easily verified by multi-agent models. If we consider the floating population to be an ER network, it is possible to explain the effect of decreasing returns to scale.

In this study, we found that the distance decay effect not only affects the scale of population flows, but also affects the transmission performance of population flow in the epidemic. It is generally believed that all the distance decay effects have been taken into account in population mobility, but during this study we found that the epidemic transmission at different distances was still different, even with the same seed number (number of infected cases) and population mobility scale. Generally, the farther the distance is, the worse the diffusion ability is; namely, there is an additional distance decay effect. The authors believe that the further existence of a distance decay effect is due to the different modes of long-distance evaluation; L.M. revised the manuscript and prepared part of data; T.P. processed and analyzed the population mobility data; S.W. and C.X. participated in interpretation of the result and provided feedback for this manuscript. All authors read and approved the final manuscript.

Declaration of competing interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is supported by National Natural Science Foundation of China (No. 41971202, No. 41722103).

References

Aagard-Hansen, J., Nombela, N., & Alvar, J. (2010). Population movement: A key factor in the epidemiology of neglected tropical diseases: The Global epidemiology and mobility computational model. Journal of Computational Science, 3(3), 132–145. https://doi.org/10.1016/j.jocs.2010.07.002.

Brockmann, D., & Helbing, D. (2013). The hidden geometry of complex, network-driven contagion phenomena. Science, 342(6144), 1357–1342. https://doi.org/10.1126/science.1245200.

Chiunazi, M., Davis, J. T., Ajelli, M., Giovinetti, C., & Vespignani, A. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science, 368(6489), 395–400. https://doi.org/10.1126/science.aba9757.

Dai, D. (2011). Racial/ethnic and socioeconomic disparities in urban green space accessibility: Where to intervene? Landscape and Urban Planning, 102(4), 234–244. https://doi.org/10.1016/j.landurbplan.2011.05.002.

Eilertsen, L. J. (2007). NetLogo, a multi-agent simulation environment. Artificial Life, 12 (3), 303–311. https://doi.org/10.1162/107676207782123785.

Erdős, P., & Rényi, A. (1961). On the strength of connectedness of a random graph. Acta Mathematica Hungarica, 12(1), 261–267. https://doi.org/10.1007/BF02066689.

Eubank, S., Goh, K. I., Kumar, V. S. A., Marathe, M. V., Minh, V. N., Clipazzi, L., & Wang, N. (2004). Modelling disease outbreaks in realistic urban social networks. Nature, 429, 180–184. https://doi.org/10.1038/nature02541.

Ferguson, N. M., Cummings, D. A. T., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., … Burke, D. S. (2005). Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature, 437, 209–214. https://doi.org/10.1038/nature04017.

Gao, P., Zhang, H., Wu, Z., & Wang, J. (2020). Visualising the expansion and spread of coronavirus disease 2019 by cartograms. Environment and Planning A, 52(4), 608–701. https://doi.org/10.1177/0144122120910162.

Gu, C., Zhu, J., Sun, Y., Zhou, K., & Gu, J. (2020). The infection point about COVID-19 may have passed. Science Bulletin, 65(11), 865–867. https://doi.org/10.1016/j.scib.2020.02.025.

Hu, H. (1955). The distribution of population in China, with statistics and maps. Acta Geographica Sinica, 02, 191–192 (In Chinese).

Jia, J. S., Lu, X., Yuan, Y., Xu, G., Jin, J., & Christakis, N. A. (2020). Population flow drives spatio-temporal distribution of COVID-19 in China. Nature, 582, 389–394. https://doi.org/10.1038/s41586-020-2284-y.

Jiao, J., Wang, J., & Jin, F. (2017). Impacts of high-speed rail lines on the city network in China. Journal of Transport Geography, 60, 257–266. https://doi.org/10.1016/j.jtrangeo.2017.03.010.

Jiao, J., Wang, J., Jin, F., & Dunford, M. (2014). Impacts on accessibility of China’s present and future HSR network. Journal of Transport Geography, 40, 123–132. https://doi.org/10.1016/j.jtrangeo.2014.07.004.

Kamien, W. O., & Mckendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. Proceedings of The Royal Society A Mathematical Physical and Engineering, 115(772), 701–721. https://doi.org/10.1098/rspa.1927.0118.

Khan, A., Arino, J., Hu, W., Raposo, P., Sears, J., Calderon, F., … Gardam, M. (2009). Spread of a novel influenza A (H1N1) virus via global airline transportation. The New England Journal of Medicine, 361(2), 212–214. https://doi.org/10.1056/NEJMCM0904559.

Kraemer, M. U. G., Yang, C., Gutierrez, B., Wu, C., Klein, B., Klein, B., … Scarpino, S. M. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. Science, 368(6493), 493–497. https://doi.org/10.1126/science.abb4218.

Lai, P., Chow, C., Wong, H., Kwong, K., Kwan, Y., Liu, S., … Wong, W. (2015). An early warning system for detecting H1N1 disease outbreak – A spatio-temporal approach. International Journal of Geographical Information Science, 29(7), 1251–1268. https://doi.org/10.1080/13658816.2015.1030671.

Lai, S., Ruktanonchai, N. W., Zhou, L., Prosper, O., Luo, W., Floyd, J. R., … Tatem, A. J. (2020). Effect of non-pharmaceutical interventions to contain COVID-19 in China. Nature. https://doi.org/10.1038/s41586-020-2293-x. Published online: 4 May 2020.

Murray, J. D. (1993). Mathematical biology (2nd ed.). Berlin Heidelberg: Springer-Verlag.

Paine, S., Mercer, G. N., Kelly, P. M., Bandanaanayake, D., Baker, M. G., Huang, Q. S., … Hope, V. (2010). Transmissibility of 2009 pandemic influenza A (H1N1) in New Zealand: Effective reproduction number and influence of age and importations. Euro Surveillance, 15(24), 9–17. https://doi.org/10.1185/1471234-10-176.
Pan, J., & Lai, J. (2019). Spatial pattern of population mobility among cities in China: Case study of the national day plus mid-autumn festival based on Tencent migration data. Cities, 94, 55–69. https://doi.org/10.1016/j.cities.2019.05.022.

Prothero, R. M. (1963). Population mobility and trypanosomiasis in Africa. Bulletin of the World Health Organization, 28(5–6), 615–626.

Read, J. M., Bridgen, J. R. E., Cummings, D. A. T., Ho, A., & Jewell, C. P. (2020). Novel coronavirus 2019-nCoV: Early estimation of epidemiological parameters and epidemic predictions. medRxiv. https://doi.org/10.1101/2020.01.23.20018549, published online Jan 24. (preprint).

Schlagenhauf, P. P., Weld, L., Goorhuis, A., Gautret, P., Weber, R., Sonnenburg, F., … Grobusch, M. P. (2014). Travel-associated infection presenting in Europe (2008–12): An analysis of EuroTravNet longitudinal, surveillance data, and evaluation of the effect of the pre-travel consultation. The Lancet Infectious Diseases, 15(3), 263. https://doi.org/10.1016/S1473-3099(14)71000-X.

Shi, Q., & Liu, T. (2020). Should internal migrants be held accountable for spreading COVID-19? Environment and Planning A, 52(4), 695-697. https://doi.org/10.1177/0308518X20916764.

Small, M., & Tse, C. K. (2005). Small world and scale free model of transmission of SARS. International Journal of Bifurcation and Chaos, 15, 1745–1755. https://doi.org/10.1142/S0218127405012776.

Tang, M., Liu, Z., & Li, B. (2009). Epidemic spreading by objective traveling. Europhysics Letters, 87(1), 18005. https://doi.org/10.1209/0295-5075/87/18005.

Tizzoni, M., Bajardi, P., Poletto, C., Ramasco, J. J., Balcan, D., Gonçalves, B. L., … Vespiagnani, A. (2012). Real-time numerical forecast of global epidemic spreading: Case study of 2009 A/H1N1pdm. BMC Medicine, 10(1), 165. https://doi.org/10.1186/1741-7015-10-165.

Wang, L., & Li, X. (2014). Spatial epidemiology of networked metapopulation: An overview. Chinese Science Bulletin, 59(28), 3511–3522. https://doi.org/10.1007/s11434-014-0499-8.

Y. Wei et al. Cities 110 (2021) 103010