Transportation Safety and Environment, 2021, Vol. 3, No. 1 21–35

doi: 10.1093/tse/tdaa033
Advance Access publication 2021 January 18
Research Article

RESEARCH ARTICLE

Correlation of the epidemic spread of COVID-19 and urban population migration in the major cities of Hubei Province, China

Qun Chen[1], Jiao Yan[1,*], Helai Huang[1] and Xi Zhang[2]

[1]Central South University, School of Traffic and Transportation Engineering, Changsha, 410075 Hunan, China and [2]Central South University, School of Materials Science and Engineering, Changsha, 410075 Hunan, China

[*]Corresponding author. E-mail: 925544687@qq.com

Abstract

This study analyses the relationship between the epidemic spread of COVID-19 and urban population migration in Hubei Province, China. Based on an improved gravity model, the population inflow numbers for each city from 10 January to 23 February are estimated. A correlation analysis is done to reveal the impact of population inflow on the number of infected people in the 14 days after 23 January, the day Wuhan was locked down. The results show that: (i) the population outflow from Wuhan was mostly distributed between Xiaogan, Huanggang, Ezhou and Huangshi in Hubei Province; (ii) the number of accumulated confirmed patients is closely associated with inflows from Wuhan, which displayed by correlation coefficient 1 with a mean of 0.88 and a maximum of 0.93. Meanwhile, there is a weak correlation between the number of people that came from cities except Wuhan and accumulated confirmed patients, which indicated by correlation coefficient 2 with a mean of 0.65 and a maximum of 0.75; and (iii) the total population inflow is a greater predictor of epidemic spread than the population inflow from Wuhan.

Keywords: population migration; correlation analysis; gravity model; COVID-19; Hubei Province

1 Introduction

At the end of 2019, coronavirus disease 2019 (COVID-19) was first discovered in China and rapidly reported by various overseas regions. As one of the early disaster areas of the COVID-19, Wuhan, is located at the hub of the transportation network, and the outbreak coincided with the Spring Festival travel rush (also known

Received: 23 June 2020; **Revised:** 6 October 2020; **Accepted:** 12 November 2020

as *Chunyun*), the annual period of mass migration, large-scale population movement became a booster for the spread of the disease [1, 2]. COVID-19, a new respiratory infection, showed a high correlation with population migration [3, 4], and their relationship is mainly reflected in the incidence of COVID-19 in cities adjacent to Wuhan, Guangzhou, Shenzhen, Shanghai, Changsha and so on, which was significantly higher than the national average. These cities had conducted high-intensity migration movements with Wuhan before the Spring Festival in 2020, according to Baidu migration data. Consequently, it is necessary to conduct research on the mechanism of population migration and epidemic transmission when studying the transmission dynamics and spatial distribution characteristics of COVID-19.

The earliest mathematical model of population migration is the classic gravity model. Zipf formally proposed the gravity model based on the research of British demographer Lewinstein [5]. Subsequently, the model was widely applied to the study of population migration [6, 7]. However, due to the major shortcomings of the classic gravity model, such as the loss of characteristics related to individual movement and its poor adaptability to different spatial scales, many scholars began more sophisticated theoretical research after the 1950s.

Migration theories can be divided into micro-level, meso-level and macro-level theories. Macro-theories of migration regard migration as a social phenomenon and usually explain migration as geographical differences in the supply and demand of labour or income differentials [8, 9]. The typical representatives of macro theories of migration are neoclassical macro-migration theory and dual labour-market theory. But some scholars also take the political setting as an important factor in migration [10]. DaVanzo proposed a life-cycle model of population migration after studying mobility in Philadelphia [11]. A common shortcoming of macro-level models is their overly narrow focus, with only one specific pull factor and with no deeper analysis of migrant decision-making [12], and their failure to consider different types of migration and individual differences. Micro theories of migration focus on the study of migration motivation and decision-making on the individual or family level. Lee provided a push–pull framework to predict the volume of immigration under various conditions [13]. Sjaastad proposed the human-capital approach based on individual investment decisions and revealed that people migrate when the positive result is expected [14]. Crawford derived a formula to calculate the expected benefits and migration costs and introduced some non-economic factors into the value-expectancy model [15]. Compared to individual behavioural decision-making models, the New Economics of Labour Migration weakens the impact of income differences and emphasizes the significance of remittances from migrants [16]. At the end of the 20th century, the immigrant-network theory was born. Migrant networks, combinations of a series of relationships, decrease the costs and risks of migration and thus attract more migrants, and the new migrants will in turn expand the migration networks [17]. Moreover, some scholars have associated cross-border migration with economic globalization and global capital flow and have established the world-system theory [18, 19].

Due to the limited power of individual characteristics or macroeconomic or social factors to explain population migration, some scholars tend to combine macro theories with micro theories and census data to conduct research. In view of the 1988 Chinese fertility-sampling survey data, Liang and White found that the young adults and the more educated dominated the migration stream with a balanced sex ratio between 1983 and 1988, and that marital status and sex have no significant effect on migration [20]. Duan's research showed that men are more prone to interprovincial migration than women [21]. However, Tang and Ma found that more females in their early twenties were on the move than males in the late 1990s, based on microdata from the 2000 Chinese census [22]. Yuan et al. applied a multi-level logit model to analyse the choice of migration destination in large cities [23]. Fan pointed out that migration is related to regional development [24]. Gonzáles et al. studied the trajectory of 100 000 anonymized mobile phone users to determine a general law of human mobility [25]. Zhang and Cen used the ordinary multivariate-regression and spatial-regression models to analyse intra- and interprovincial population migration, respectively [26]. Using the Baidu Migration big data, Jiang and Wang measured the characteristics of an urban population-mobility network covering 334 cities [27].

With the unprecedented increase in population migration worldwide, the existence of a mass floating population has not only promoted the development of the economy and society, but also caused a series of social problems, among which

the most concerning is the spread of diseases. Historically, population movement has contributed to disease transmission [28, 29]. Neouimine studied the causes of increased incidence of leishmaniasis in some member states of the Eastern Mediterranean Region, and found that the influx of a non-immune population into natural foci of transmission is a very important factor [30]. Gushulak and MacPherson found that some groups of migrants and mobile populations will face a higher risk of certain infections [31]. Garnett and Lewis proposed that population migration will increase the contact rate in the SIR (susceptible-infected-recovered) model [32]. Gani et al. built a variety of models of the spread and quarantine of HIV infection in a prison system by assuming a constant flow into and out of the population and a constant total population size [33]. Brauer and Driessche constructed an SIS (susceptible-infected-susceptible) model that included the immigration of infective individuals and variable population size [34]. Wang and Mulone established an epidemic model to explain the dynamics of disease spread between two areas caused by population dispersal [35]. Takeuchi et al. proved that travel restrictions for infected individuals are important for controlling the spread of disease [36].

Recently, some scholars have introduced migration into the infectious disease dynamics models to study the impact of migration on COVID-19 transmission. Wu et al. used the SEIR (susceptible-exposed-infectious-removed) model to predict the COVID-19 infection cases imported from Wuhan to five cities in China [37]. Based on population migration before and after the quarantine, the team of Zhong Nanshan used the SEIR model and a machine-learning artificial-intelligence approach to predict that the epidemic in China would peak by late February [2]. Tang et al. assessed the impact of public-health interventions on infection based on a deterministic compartmental model and emphasized the importance of strict isolation [38]. Read ang Bridgen estimated the number of immigrants using flight information, and obtained the number of infected people in Wuhan and other parts of the country as of 22 January [39]. Tang et al. examined the impact of migration patterns and the proportion of infected people within the imported population on the risk of secondary outbreaks based on a novel stochastic discrete transmission model for COVID-19 [40]. Xu et al. provide a statistical analysis of the destination distribution of people leaving Wuhan, and studied the influence of population outflow on epidemic spread [41]. Liu et al. studied the spatio-temporal spreading characteristics of COVID-19 at the county spatial scale, and further evaluated the risk owing to population movement after the Spring Festival [42].

The existing literature in this field, however, has focused on long-term population migration centred on the household-registration system. The intensity of migration is usually equal to the number of people moving in or out, or the size of the urban floating population [43]. These models also ignore cross-regional migration movements for non-residence purposes (such as work, study, holiday home stays) and short-term migration movements (such as tourism, visiting relatives), which are closely related to epidemic transmission [44]. Another limitation lies in the fact that this work overlooks the differences between inter- and intra-provincial human mobility. Scholars have proved that inter- and intra-provincial floating populations are two groups with significant differences [45], but few studies have distinguished between or compared them. Recent evidence has proved the significant effect of spatial scales on human mobility, raising the need for formulating a universal model suited to human mobility at different levels and spatial scales [46]. In addition, due to the fact that the COVID-19 outbreak in China coincided with *Chunyun*, the total population and the population inflow and outflow changed rapidly and dynamically. In addition, COVID-19 has an incubation period and unobvious early-onset symptoms. Despite their explanatory power, existing infectious disease dynamics models still have certain limitations on the simulation and prediction of epidemic transmission in the first stage of free migration.

In response to the above two limitations, our study draws on Fan and Yan's argument and extends it to the gravity model. Recent studies on the dynamic model of COVID-19 have suggested the complementarity of population migration, statistical data, control measures, clinicopathologic features and infectious disease dynamics theories in explaining epidemic transmission [2, 47]. Following their conclusions, we develop an improved gravitational model to explore the interrelation between population migration and the spread of COVID-19 in Hubei Province.

This study extends research on population migration and infectious disease in two major dimensions. First, it is among the first

attempts to analyse short-term migration within a province based on an improved gravitational model. Despite the long-standing research on population flow, the literature is still lacking in the calculation of short-term migration and the holistic analysis of migration networks within an area. By introducing the direction coefficient and traffic accessibility, this model achieves a better understanding of the choice of destination and seasonal fluctuation of human mobility. Second, in light of the mass migration before the strict epidemic prevention and control measures, we conduct an analysis of the correlation between population flow and the spread of COVID-19 and put forward interesting insights into the impact of population inflow from Wuhan and the total population inflow, which provide a method for predicting the geographical distribution of infectious diseases at an early stage.

2 Methods

The first application of the gravity model in population studies was the analysis of population migration in Britain by the demographer Lewinstein in 1880 [48]. The original gravity model included only two variables: population size and spatial distance. Many scholars subsequently expanded the model, and factors such as education, age, income, politics and so on have been gradually incorporated into the model to enhance its explanatory power for the study of migration [49–51].

The definition of the gravity model is similar to the mechanism of human mobility. Compared with the micro model, the gravity model shows less dependence on data, with a simpler model-parameter calibration. This model is a comprehensive consideration of both moving into and moving out of an area, rather than only one of these. The migration intensity calculated by the model represents the total potential of human mobility between cities without direction. The directional coefficient needs to be determined according to the trait of migration at different stages. Compared to the traditional gravity model, the improved gravity model can reflect the dynamic mechanism of population flow more objectively and accurately, with the inclusion of transportation accessibility, employment opportunities and public services, and the seasonal law of the fluid population. The difference in demographic characteristics between non-provincial cities is relatively small, and the increase in the proportion of urban-to-urban migration has led to the insignificant impact of factors such as gender, marital status and household registration type on the probability of intra-provincial migration [23, 45]. Taking all the above-mentioned into account, we opt for an improved gravity model for this study.

Human mobility results from the joint effects of attraction and resistance. Existing studies have proved that the factors associated with attraction of migration include population size, socioeconomic development level, public services and resources, and distance between cities [26, 52, 53]. Cities with high levels of economic development and large populations are more attractive to immigrants, and regions with more employment opportunities also have greater competitiveness. The resistance to population movement comes mainly from spatial distance; the greater the distance between cities, the greater the increase in mobility costs [54].

Traffic is also an important factor affecting population mobility. There is a significant correlation between transportation networks and human mobility. The improvement of transportation network can greatly improve the ability of migration and, consequently, promote human mobility [55, 56].

The factors related to population migration are shown in Fig. 1. In this study, a gravity model describing the spatial interaction is used to predict migration between cities. The gravity model is defined as follows:

$$M_{ij} = A_{ij} \frac{Q_i \cdot Q_j}{d_{ij}^{\varepsilon}} \quad (1)$$

where M_{ij} represents the intensity of population migration between city i and city j; Q_i and Q_j respectively indicate the quality of city i and city j; A_{ij} is the traffic accessibility between city i and city j, which is a specific value of the comprehensive traffic accessibility; d_{ij} shows the spatial distance between city i and city j; and ε is the distance-attenuation coefficient.

The distance-attenuation coefficient reflects the sensitivity of migration to spatial distance, and is generally taken as 1 or 2 [54, 57]. Migration shows considerable power to conquer distance during the Spring Festival travel rush, as past experience has proved. Therefore, this study sets the distance-attenuation coefficient as 1 to fit the low sensitivity of human mobility to distance.

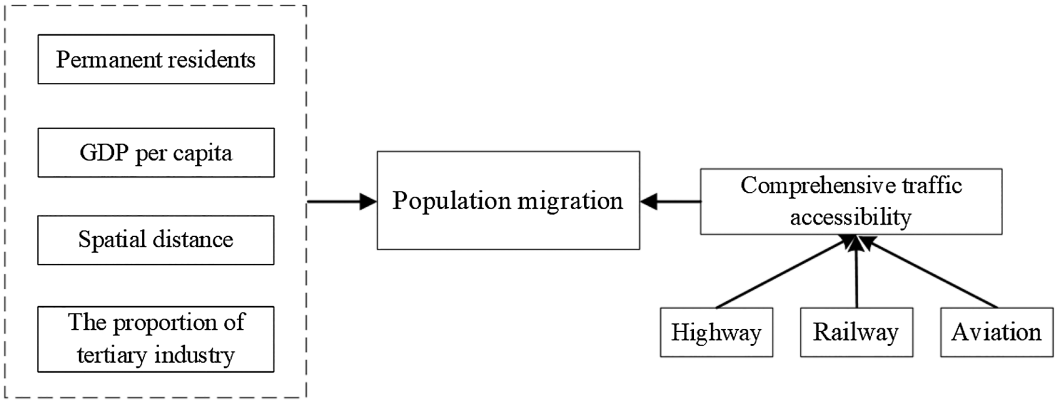

Fig. 1. Factors influencing population migration

Population flow as calculated by this model conveys the capability of human mobility without directionality. Nonetheless, population migration in different periods expresses obvious directionality. For instance, the floating population before the Spring Festival is composed mainly of returnees, showing characteristics of moving from developed areas to underdeveloped areas. In contrast, after the Spring Festival, population migration shows characteristics of moving from underdeveloped regions to developed regions for the resumption of work and school. Hence, this study defines M'_{ij} as movement from city i to city j, and M''_{ij} as that from city j to city i, and Equation (2) is obtained.

$$M'_{ij} + M''_{ij} = M_{ij}$$
$$M'_{ij} : M''_{ij} = a \quad (2)$$

where a is direction coefficient.

The specific implementation process of the model is shown in Fig. 2.

2.1 City quality

City quality is a comprehensive manifestation of a city's urban population, socio-economic level and employment opportunities, and constitutes a positive correlation with these. The existing literature has proved that socio-economic level, measured by GDP per capita, has a significant impact on population mobility [58]. In addition, the proportion of tertiary industry reflects the employment opportunities and public facilities of the city to a certain extent. For these reasons, the formula for calculating city quality is as follows:

$$Q_i = P_i \cdot G_i \cdot T_i \quad (3)$$

where P_i refers to the number of permanent residents in city i, G_i delineates the GDP per capita in city i, and T_i presents the proportion of tertiary industry in city i.

2.2 Comprehensive traffic accessibility between cities

Traffic accessibility refers to the convenience of and lack of impediment to transportation between two cities, which is the combined effect of different traffic modes. In view of the low proportion of water transport, this study analyses only three traffic modes: road, rail and aviation. By means of the topology-analysis method, their accessibility is obtained separately, and thus the comprehensive traffic accessibility is calculated by weighted superposition. The specific formula for complete traffic accessibility is as follows:

$$A_{ij} = \alpha \cdot h_{ij} + \beta \cdot r_{ij} + \gamma \cdot a_{ij} \quad (4)$$

where h_{ij}, r_{ij} and a_{ij} respectively delineate the accessibility of the road, rail and civil aviation networks between city i and city j; and α, β and γ are the weight coefficients of road travel, rail travel and civil aviation in the regional passenger-transport network, where $\alpha + \beta + \gamma = 1$. Taking road transportation as an example, the calculation of road network accessibility is done as follows:

(i) Sketch the topology map of the road network:
Cross-city passenger transport is achieved mostly by expressways connecting cities; consequently, only the expressway is premeditated in the road passenger transport network. First, we construct a regional road (expressway) network topology map based on the distribution of road (expressway) routes, and mark the number when the route

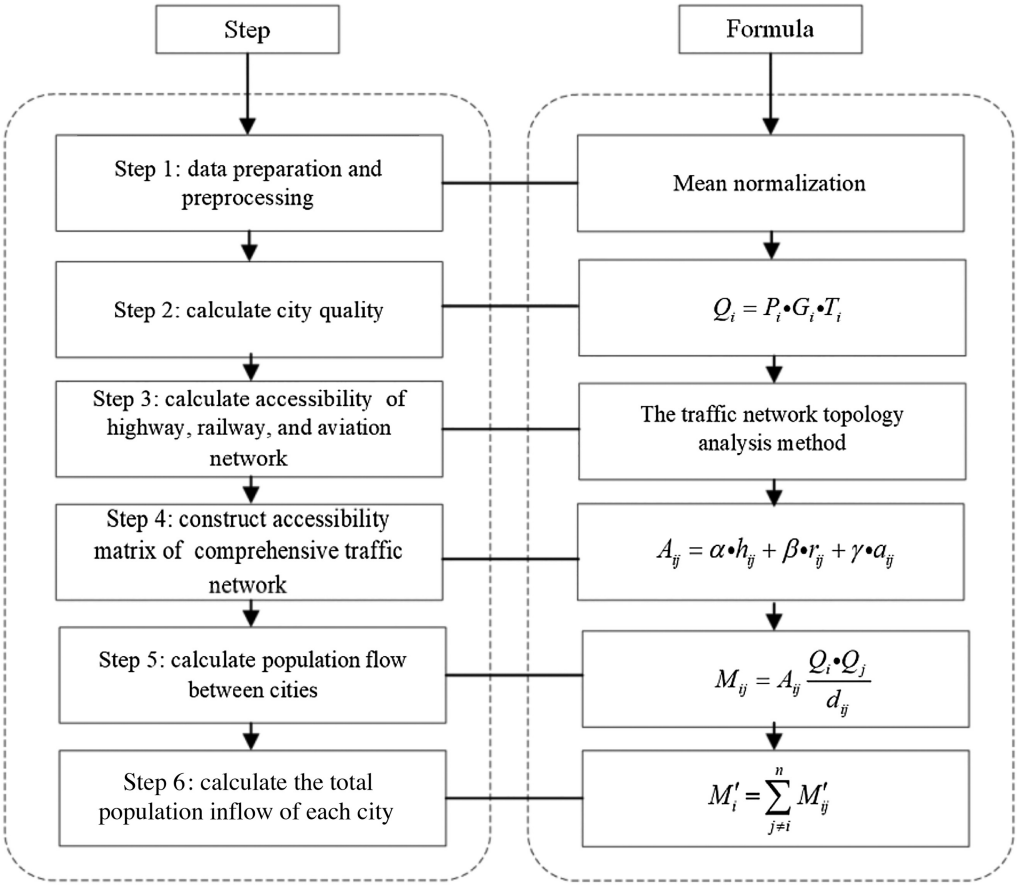

Fig. 2. Implementation of the improved gravity model

between adjacent cities exceeds one. This study assumes that there is no difference in the transport capacity of different road (expressway) routes.

(ii) Construct road-network accessibility matrix *H*:

Taking the city as a node, a node-connection matrix (road network-accessibility matrix) is constructed, where h_{ij} represents the accessibility of the road network between city *i* and city *j*. The formula for calculating network accessibility is as follows:

$$h_{ij} = 1/h'_{ij}$$
$$h'_{ij} = \begin{cases} 1/n^h_{ij} & \text{city i and city } j \text{ are adjacent} \\ \min\{h'_{ir} + h'_{rj}\}, & \text{city i and city } j \text{ are not adjacent} \end{cases} \quad (5)$$

where n^h_{ij} represents the number of roads connecting city *i* and city *j*, *r* stands for the central city between them, and h'_{ij} represents the obstruction coefficient of road transport between them. The larger n^h_{ij} is, the larger h_{ij} is, and the better the accessibility of the road network is.

(iii) Define the traffic-mode split rate:

Given that demand for transportation services seriously outstrips supply during the Spring Festival travel rush, the differences between travel modes are ignored to the greatest extent. In other words, the traffic-mode split rate is determined by each mode's transportation capacity.

2.3 Urban population inflow and outflow

Assuming that there are *n* cities in the study region, the formulas for calculating the total outflows and total inflows of a city are as shown in Equations (6) and (7):

$$M'_i = \sum_{j \neq i}^{n} M'_{ij}, i, j = 1, 2, \ldots n \quad (6)$$

$$M''_i = \sum_{j \neq i}^{n} M''_{ij}, i, j = 1, 2, \ldots n \quad (7)$$

Table 1. Quality of each city in Hubei Province

City	Wuhan	Xiaogan	Huanggang	Jingzhou	Ezhou
Code	1	2	3	4	5
Quality	9.40	0.81	0.96	0.89	0.45
City	Suizhou	Xiangyang	Huangshi	Jingmen	Yichang
Code	6	7	8	9	10
Quality	0.44	1.94	0.65	0.79	1.79
City	Xianning	Shiyan	Xiantao	Tianmen	Enshi
Code	11	12	13	14	15
Quality	0.59	0.87	0.35	0.24	0.47
City	QianJiang	Shennongjia			
Code	16	17			
Quality	0.33	0.02			

where M_i' represents the number of outflows from city i and M_i'' indicates the number of inflows into city i.

3 Hubei Province as a case study

Since the outbreak of COVID-19 in late December 2019, 66337 confirmed cases in total had been reported in Hubei as of 24:00 on 28 February 2020. This section will estimate human migration between subordinate cities in Hubei Province and analyse the relationship between the spread of COVID-19 and population migration.

3.1 Study region and data preparation

Our research scope covers 17 cities in Hubei Province, including 12 prefecture-level cities, one autonomous prefecture, three provincial-controlled divisions, and one provincial-controlled forest area. The information sources include: (i) territory, population, economic and industrial structure data from the Statistical Yearbook published on the official website of the Hubei Provincial Bureau of Statistics, which includes statistics up to the beginning of 2019; (ii) the linear distance between cities, calculated based on longitude and latitude data for each city centre provided by Baidu Maps; and (iii) administrative division and transportation network maps from the Hubei Natural Resources Department.

3.2 Calculation of city quality

Table 1 shows the city quality calculated using Equation (3). Unless otherwise specified, all data in this paper refers to the results after mean normalization, and the dimension of each parameter is 1, without labelling.

3.3 Calculation of distance between cities

From the longitude and latitude information for each city centre, the distance between cities, d_{ij}, is obtained, as shown in Table 2.

$$\bar{d}_{ij} = d_{ij} \cdot \frac{n(n-1)}{2\sum_i^n \sum_j^n d_{ij}}, \quad n = 17, i, j = 1, 2, \ldots 17 \quad (8)$$

In Equation (8), $\bar{d}_{ij}$ represents the result of standardized d_{ij}.

3.4 Construction of accessibility matrix of a regional comprehensive traffic network

Step 1: Figs 3 and 4 provide the geometric diagrams of the rail and road networks, respectively. In addition, there are three non-stop flights in Hubei Province: Wuhan–Shiyan, Wuhan–Shennongjia and Wuhan–Enshi. The numbers in these figures indicate the number of roads (rail links) between adjacent cities.

Step 2: Accessibility matrixes corresponding to different traffic are generated using Equation (5).

Step 3: The traffic-mode split rate is determined.

In the past 10 years, road, rail and air travel have shown a synchronous development trend, with each retaining a stable proportion of the passenger-transportation network, and thus their portion of passenger capacity as of 2018 (shown in Table 3) is adopted as the traffic-mode split rate.

Step 4: Three traffic-accessibility matrixes are weighted and superimposed to obtain the regional comprehensive traffic-accessibility matrix H, as shown in Table 4.

3.5 Calculation of the intensity of population migration

The intensity matrix of population migration between cities is calculated using Equation (1), and the results are shown in Table 5.

4 Correlation analysis of the epidemic spread of COVID-19 and human mobility

After the lethal case of COVID-19 reported in Wuhan on 9 January 2020, the Spring Festival travel rush began. During the Spring Festival travel rush—10 to 23 January—a great number of people moved freely between cities. To contain the spread of COVID-19, cities in Hubei Province

Table 2. Matrix of standard distances between cities

City	City 1	City 2	City 3	City 4	City 5	City 6	City 7	City 8	City 9	City 10	City 11	City 12	City 13	City 14	City 15	City 16	City 17
City 1	–	0.24	0.27	0.92	0.27	0.71	1.20	0.37	0.94	1.31	0.36	1.84	0.43	0.52	2.10	0.69	1.74
City 2	0.24	–	0.49	0.80	0.50	0.47	0.96	0.61	0.75	1.15	0.55	1.61	0.40	0.38	1.95	0.59	1.53
City 3	0.27	0.49	–	1.17	0.03	0.93	1.45	0.13	1.21	1.58	0.39	2.09	0.67	0.78	2.36	0.94	2.01
City 4	0.92	0.80	1.17	–	1.16	0.86	0.89	1.23	0.36	0.45	0.93	1.32	0.50	0.42	1.19	0.23	1.00
City 5	0.27	0.50	0.03	1.16	–	0.94	1.46	0.11	1.21	1.58	0.37	2.11	0.66	0.78	2.35	0.93	2.01
City 6	0.71	0.47	0.93	0.86	0.94	–	0.55	1.05	0.61	1.03	1.01	1.20	0.72	0.55	1.82	0.75	1.25
City 7	1.20	0.96	1.45	0.89	1.46	0.55	–	1.57	0.53	0.79	1.45	0.65	1.05	0.84	1.47	0.92	0.77
City 8	0.37	0.61	0.13	1.23	0.11	1.05	1.57	–	1.31	1.66	0.37	2.21	0.73	0.87	2.42	1.00	2.11
City 9	0.94	0.75	1.21	0.36	1.21	0.61	0.53	1.31	–	0.42	1.08	1.01	0.64	0.44	1.24	0.43	0.80
City 10	1.31	1.15	1.58	0.45	1.58	1.03	0.79	1.66	0.42	–	1.38	0.99	0.93	0.79	0.81	0.67	0.57
City 11	0.36	0.55	0.39	0.93	0.37	1.01	1.45	0.37	1.08	1.38	–	2.06	0.45	0.65	2.11	0.71	1.88
City 12	1.84	1.61	2.09	1.32	2.11	1.20	0.65	2.21	1.01	0.99	2.06	–	1.63	1.42	1.30	1.44	0.55
City 13	0.43	0.40	0.67	0.50	0.66	0.72	1.05	0.73	0.64	0.93	0.45	1.63	–	0.21	1.69	0.27	1.42
City 14	0.52	0.38	0.78	0.42	0.78	0.55	0.84	0.87	0.44	0.79	0.65	1.42	0.21	–	1.59	0.22	1.24
City 15	2.10	1.95	2.36	1.19	2.35	1.82	1.47	2.42	1.24	0.81	2.11	1.30	1.69	1.59	–	1.42	0.78
City 16	0.69	0.59	0.94	0.23	0.93	0.75	0.92	1.00	0.43	0.67	0.71	1.44	0.27	0.22	1.42	–	1.18
City 17	1.74	1.53	2.01	1.00	2.01	1.25	0.77	2.11	0.80	0.57	1.88	0.55	1.42	1.24	0.78	1.18	–

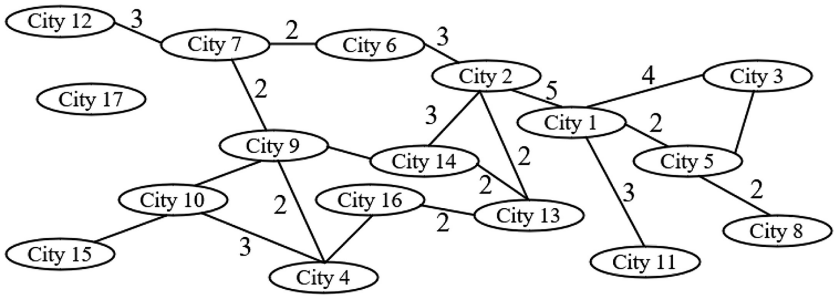

Fig. 3. Topology of Hubei Province rail network

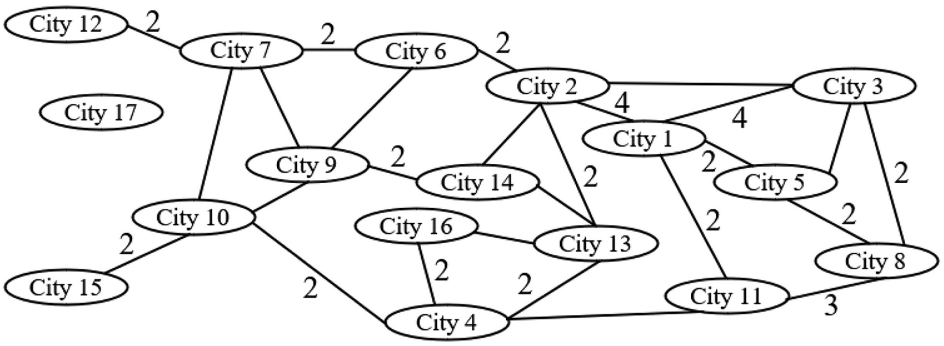

Fig. 4. Topology of Hubei Province road network

Table 3. Structure of Hubei Province passenger-transportation network, 2018

Traffic mode	Road	Rail	Civil aviation
Proportion of passenger capacity (%)	82.58	16.01	1.41

successively announced lockdowns from 23 January onwards, the date on which Wuhan issued a lockdown notice. Population migration between and within cities was then strictly controlled. Cases of COVID-19 infection started to appear all over the country from 17 January 2020 (date of first report) onwards, especially in Hubei Province, which was the main receiving area of the population outflow from Wuhan. The mass epidemic outbreak after the Spring Festival was closely related to the population migration before the holiday. It is thus critical to assess the effect of human

Table 4. Accessibility matrix of comprehensive traffic network in Hubei Province

City	City 1	City 2	City 3	City 4	City 5	City 6	City 7	City 8	City 9	City 10	City 11	City 12	City 13	City 14	City 15	City 16	City 17
City 1	–	2.48	2.38	0.45	1.19	0.85	0.49	0.76	0.35	0.32	1.29	0.36	0.80	0.58	0.26	0.37	0.01
City 2	2.48	–	1.21	0.55	0.80	1.29	0.61	0.58	0.40	0.37	0.85	0.41	1.19	0.79	0.28	0.43	0.00
City 3	2.38	1.21	–	0.37	0.79	0.62	0.41	1.07	0.30	0.28	0.83	0.31	0.60	0.46	0.22	0.32	0.00
City 4	0.45	0.55	0.37	–	0.32	0.40	0.43	0.41	0.69	1.29	0.54	0.32	1.06	0.40	0.57	1.09	0.00
City 5	1.19	0.80	0.79	0.32	–	0.49	0.35	1.19	0.27	0.25	0.71	0.27	0.48	0.38	0.21	0.28	0.00
City 6	0.85	1.29	0.62	0.40	0.49	–	1.19	0.40	0.59	0.38	0.51	0.61	0.61	0.48	0.28	0.32	0.00
City 7	0.49	0.61	0.41	0.43	0.35	1.19	–	0.30	0.69	0.57	0.36	1.29	0.40	0.41	0.37	0.25	0.00
City 8	0.76	0.58	1.07	0.41	1.19	0.40	0.30	–	0.25	0.30	1.57	0.24	0.39	0.31	0.23	0.32	0.00
City 9	0.35	0.40	0.30	0.69	0.27	0.59	0.69	0.25	–	0.61	0.30	0.45	0.40	1.09	0.38	0.40	0.00
City 10	0.32	0.37	0.28	1.29	0.25	0.38	0.57	0.30	0.61	–	0.37	0.39	0.55	0.38	1.09	0.57	0.00
City 11	1.29	0.85	0.83	0.54	0.71	0.51	0.36	1.57	0.30	0.37	–	0.28	0.49	0.40	0.27	0.40	0.00
City 12	0.36	0.41	0.31	0.32	0.27	0.61	1.29	0.24	0.45	0.39	0.28	–	0.31	0.31	0.29	0.21	0.00
City 13	0.80	1.19	0.60	1.06	0.48	0.61	0.40	0.39	0.40	0.55	0.49	0.31	–	0.69	0.37	0.69	0.00
City 14	0.58	0.79	0.46	0.40	0.38	0.48	0.41	0.31	1.09	0.38	0.40	0.31	0.69	–	0.28	0.35	0.00
City 15	0.26	0.28	0.22	0.57	0.21	0.28	0.37	0.23	0.38	1.09	0.27	0.29	0.37	0.28	–	0.37	0.00
City 16	0.37	0.43	0.32	1.09	0.28	0.32	0.25	0.32	0.40	0.57	0.40	0.21	0.69	0.35	0.37	–	0.00
City 17	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
Total	12.92	12.24	10.18	8.88	7.99	9.03	8.12	8.31	7.19	7.74	9.15	6.06	9.03	7.31	5.49	6.35	0.01

Table 5. Matrix of population migration between cities

City	City 1	City 2	City 3	City 4	City 5	City 6	City 7	City 8	City 9	City 10	City 11	City 12	City 13	City 14	City 15	City 16	City 17
City 1	–	78.18	80.38	4.08	18.46	4.94	7.47	12.58	2.72	4.14	20.24	1.61	6.08	2.52	0.55	1.64	0.00
City 2	78.18	–	1.93	0.50	0.58	0.98	1.01	0.51	0.34	0.47	0.74	0.18	0.84	0.41	0.05	0.19	0.00
City 3	80.38	1.93	–	0.27	10.80	0.28	0.52	5.06	0.19	0.31	1.21	0.12	0.30	0.13	0.04	0.11	0.00
City 4	4.08	0.50	0.27	–	0.11	0.18	0.83	0.19	1.35	4.52	0.31	0.19	0.66	0.20	0.20	1.38	0.00
City 5	18.46	0.58	10.80	0.11	–	0.10	0.21	3.19	0.08	0.13	0.52	0.05	0.11	0.05	0.02	0.04	0.00
City 6	4.94	0.98	0.28	0.18	0.10	–	1.84	0.11	0.34	0.29	0.13	0.19	0.13	0.09	0.03	0.06	0.00
City 7	7.47	1.01	0.52	0.83	0.21	1.84	–	0.24	1.98	2.53	0.28	3.33	0.26	0.23	0.23	0.17	0.00
City 8	12.58	0.51	5.06	0.19	3.19	0.11	0.24	–	0.10	0.21	1.64	0.06	0.12	0.06	0.03	0.07	0.00
City 9	2.72	0.34	0.19	1.35	0.08	0.34	1.98	0.10	–	2.04	0.13	0.30	0.17	0.47	0.12	0.24	0.00
City 10	4.14	0.47	0.31	4.52	0.13	0.29	2.53	0.21	2.04	–	0.28	0.61	0.37	0.21	1.14	0.50	0.00
City 11	20.24	0.74	1.21	0.31	0.52	0.13	0.28	1.64	0.13	0.28	–	0.07	0.23	0.09	0.04	0.11	0.00
City 12	1.61	0.18	0.12	0.19	0.05	0.19	3.33	0.06	0.30	0.61	0.07	–	0.06	0.05	0.09	0.04	0.00
City 13	6.08	0.84	0.30	0.66	0.11	0.13	0.26	0.12	0.17	0.37	0.23	0.06	–	0.27	0.04	0.29	0.00
City 14	2.52	0.41	0.13	0.20	0.05	0.09	0.23	0.06	0.47	0.21	0.09	0.05	0.27	–	0.02	0.12	0.00
City 15	0.55	0.05	0.04	0.20	0.02	0.03	0.23	0.03	0.12	1.14	0.04	0.09	0.04	0.02	–	0.04	0.00
City 16	1.64	0.19	0.11	1.38	0.04	0.06	0.17	0.07	0.24	0.50	0.11	0.04	0.29	0.12	0.04	–	0.00
City 17	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	–
Total	245.60	86.91	101.67	14.98	34.47	9.71	21.14	24.16	10.55	17.77	26.00	6.95	9.91	4.93	2.64	5.01	0.00

mobility on the epidemic's progression. With the aim of examining the population migration during the Spring Festival travel rush, we used a modified gravity model to predict inflows and outflows. The correlation between them was studied using epidemiological data from 24 January to 8 February (the first incubation period of COVID-19). The analysis in this section is based on the following assumptions:

(i) Due to the lack of interprovincial population migration data and the particularity of Wuhan (the outbreak first reported in China), this data will be ignored in the correlation analysis.
(ii) The infected people were evenly distributed among the floating population (population inflow or outflow).
(iii) Intra-city travel intensity (related to contact rate) and secondary infection (related to contact rate) was the same in different cities; that is, if the same number of virus carriers were received, there should have the same number of infected people after the same time.
(iv) The destination choice of emigrants from Wuhan was not affected by the epidemic before the introduction of strict travel controls, which has been confirmed by some scholars [41].

The epidemic data used in this study was retrieved from the daily COVID-19 bulletin provided by the Health Commission of Hubei Province.

4.1 Analysis of urban population inflow before the Spring Festival

In consequence of the distinct characteristics of moving from developed areas to underdeveloped areas before the Spring Festival, the direction coefficient of moving from Wuhan to other cities is set as 1:9 and the opposite direction is set as 9:1, in addition, the direction coefficient of migration

Table 6. Population inflows of each city

Inflow origin	Inflows from Wuhan	Inflows from cities except Wuhan	Total population inflow
City 2	70.36	4.37	74.73
City 3	72.35	10.64	82.99
City 4	3.68	5.45	9.12
City 5	16.62	8.00	24.62
City 6	4.45	2.38	6.83
City 7	6.73	6.83	13.56
City 8	11.32	5.79	17.11
City 9	2.45	3.92	6.37
City 10	3.73	6.81	10.54
City 11	18.21	2.88	21.10
City 12	1.45	2.67	4.12
City 13	5.47	1.92	7.39
City 14	2.27	1.20	3.47
City 15	0.49	1.05	1.54
City 16	1.47	1.69	3.16
City 17	0.00	0.00	0.00
Mean	13.82	4.10	17.92

between cities except Wuhan is 1:1. The population migration data is shown in Table 6.

4.2 Relationship between urban epidemic infection and population inflow

To distinguish between inflows from Wuhan, inflows from cities except Wuhan, and the total inflows, their association with the number of infected people is discussed separately, expressed by correlation coefficient 1 ($A = 0.87$, $s = 0.04$, $R = 0.12$), correlation coefficient 2 ($A = 0.65$, $s = 0.07$, $R = 0.25$) and correlation coefficient 3 ($A = 0.88$, $s = 0.04$, $R = 0.14$). Fig. 5 shows a curve diagram describing the fluctuation of these inflows.

According to the statistics of the Chinese Spring Festival travel rush in 2020, the migration peaked on 22 January, and there was also mass migration before 24:00 on 23 January, when Wuhan was thoroughly locked down [41]. After a two-to-seven-day incubation period, virus carriers in the urban population inflow should have shown symptoms of COVID-19. Correspondingly, the correlation coefficient between the population inflow and accumulated confirmed patients should have reached a peak from 24 to 30 January. It is apparent from the graph that the correlation coefficients 1 and 3 showed a distinct growth from 26 January (the fourth day after 22 January) onwards, and stabilized at 0.9 after 28 January (the fifth day after cities were locked down), when it reached 0.9 for the first time. The mean incubation period of COVID-19 was claimed as 5.2 days by a paper published in the *New England Journal of Medicine* [47], while further study proposed a four-day median incubation period, with a five-day interquartile range (two to seven) [59]. Therefore, we consider that this trend revealed by Fig. 5 is consistent with the conclusion that the average incubation period of COVID-19 concentrates on the fourth and fifth days.

Correlation coefficient 2 started to drop from its peak of 0.75 on 1 February (the ninth day after cities were locked down) and stabilized at 0.63 around 5 February (the 13th day after the cities were closed), which is considerably different from related research results. Further analysis reveals that: (i) non-Wuhan population inflow shows the weakest explanatory power, related to its low proportion within the total population inflow, with a mean of 0.23; (ii) the three curves maintained the same rate of increase, with a stable gap between correlation coefficient 2 and correlation coefficient 1 (or 3) from 24 to 30 January. Correlation coefficient 1 (or 3) began to decline after 30 January, while correlation coefficient 2 continued to grow, narrowing the gap between them. The epidemic-prevention measures taken at different stages may explain this phenomenon: more attention was paid to the observation and quarantine of returnees among the inflow people from Wuhan at the early stage of the epidemic, with little attention paid to those from non-Wuhan cities; and (iii) to contain the spread of the epidemic, the government strengthened the management of the inflow, including of that from cities except Wuhan, and all suspected patients were put in quarantine after 2 February, prefiguring the decrease of correlation coefficient 2 at the end of the first incubation period.

A paired sample *t*-test was conducted to analyse the difference between correlation coefficients 1 and 3. The result ($t = -9.596$, $P = 0.000$) shows that there is a significant difference between the two at 0.01, which indicates that the inflow from Wuhan cannot be seen as equivalent to the total population inflow when we assess the impact of migration on epidemic transmission.

Correlation coefficient 1 still needs to be compared with correlation coefficient 3 for this reason, despite the similarity between them. We selected the infection data for 28 January, the fourth day after *Chunyun*, and 8 February, the first day after the first incubation period. The curves in Fig. 6 illustrate the epidemiological data and migration data.

As we can see from Fig. 6, the broken line representing the total population inflow is closer to that

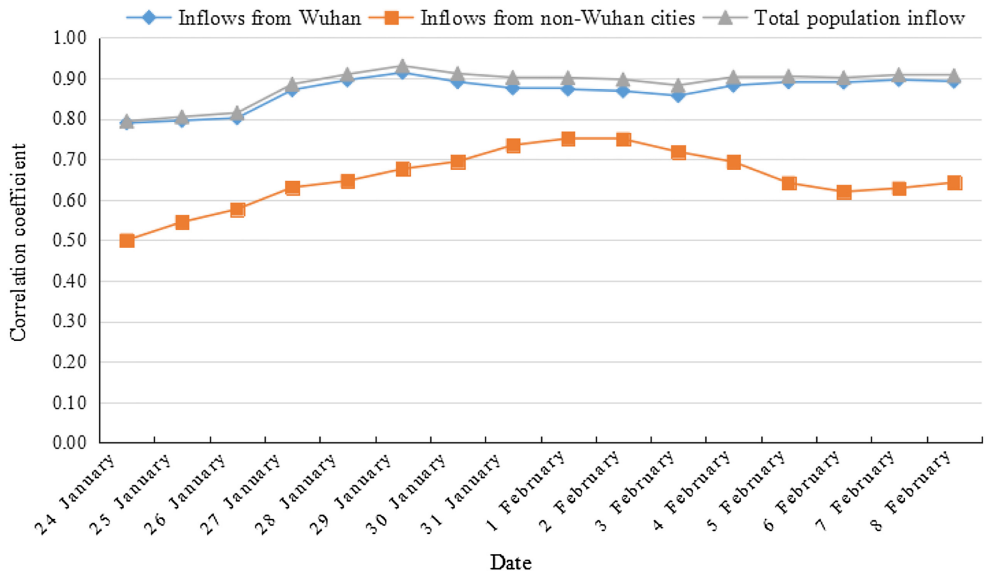

Fig. 5. Curve diagram of correlation coefficients 1, 2 and 3

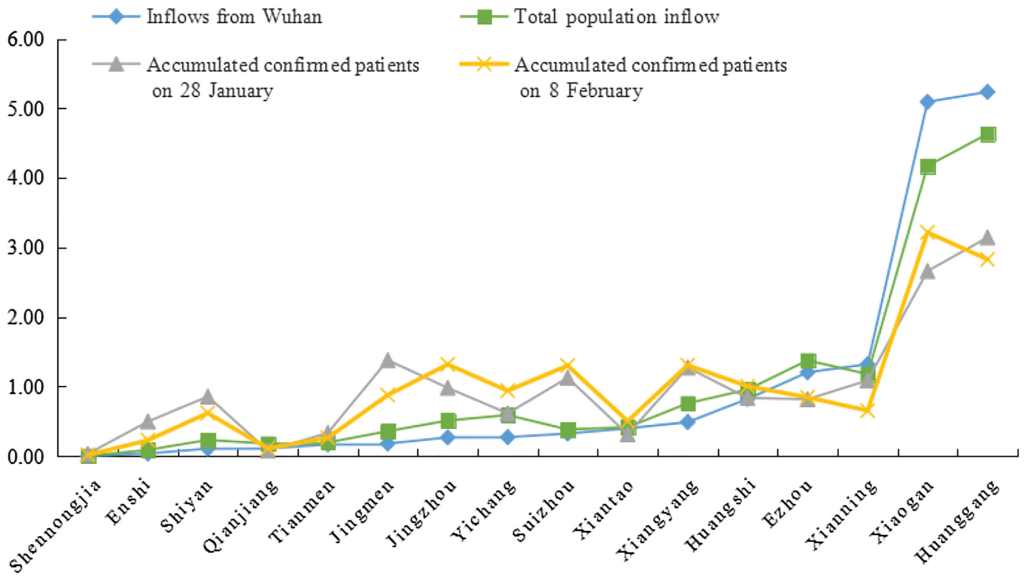

Fig. 6. Correlation between population inflow and accumulated confirmed patients

indicating the accumulated confirmed patients on 28 January and 8 February. Accordingly, we consider that the total population inflow is more relevant to the spread of the epidemic, with a higher correlation coefficient. In other words, the predictive power of the total population inflow is greater than that of the population inflow from Wuhan. And there is a positive linear correlation between the two, as shown in Fig. 7.

It is notable that overestimated results, based on the total population inflow, were obtained for Xianning, Ezhou, Qianjiang, Xiantao, Tianmen and other cities; in contrast, low predicted values appeared in Shiyan, Jingmen, Jingzhou, Suizhou and Xiangyang. We performed a geographical distribution analysis on the basis of the heat map (Fig. 8 and Fig. 9) to explain this prediction deviation. For the convenience of observation, the index of Wuhan on the heat map was set as the maximum value, which has no practical significance and cannot compare with other cities. The results show that: (i) cities with higher prediction results are closer to Wuhan, and cities with lower predicted values are located far from Wuhan. We believe that the impact of distance on population migration is overestimated in the gravity model;

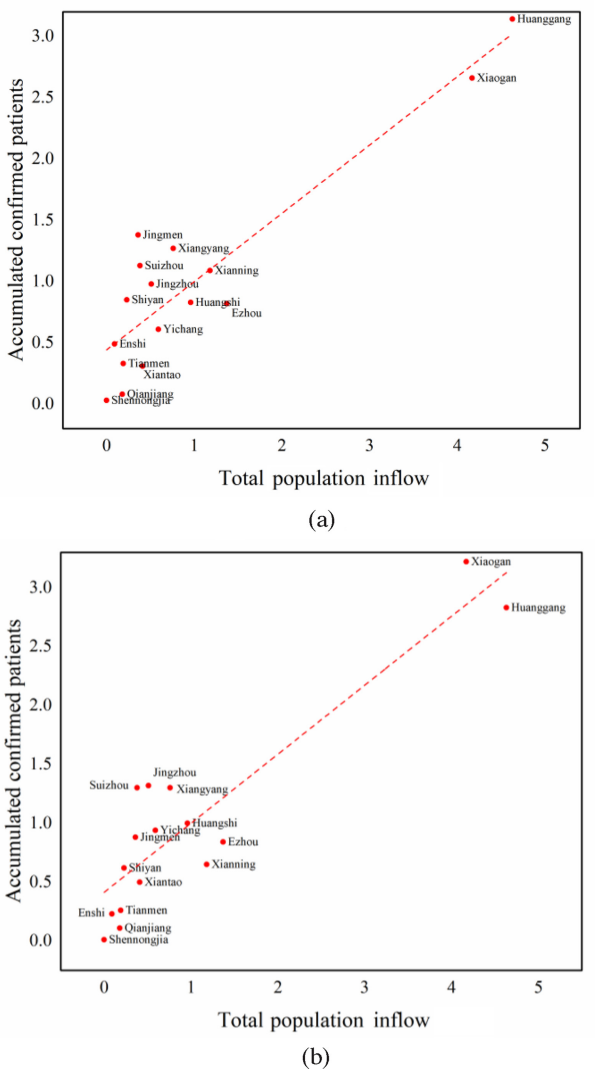

Fig. 7. Scatter plot of total population inflow and accumulated confirmed patients: (a) 28 January; (b) 8 February

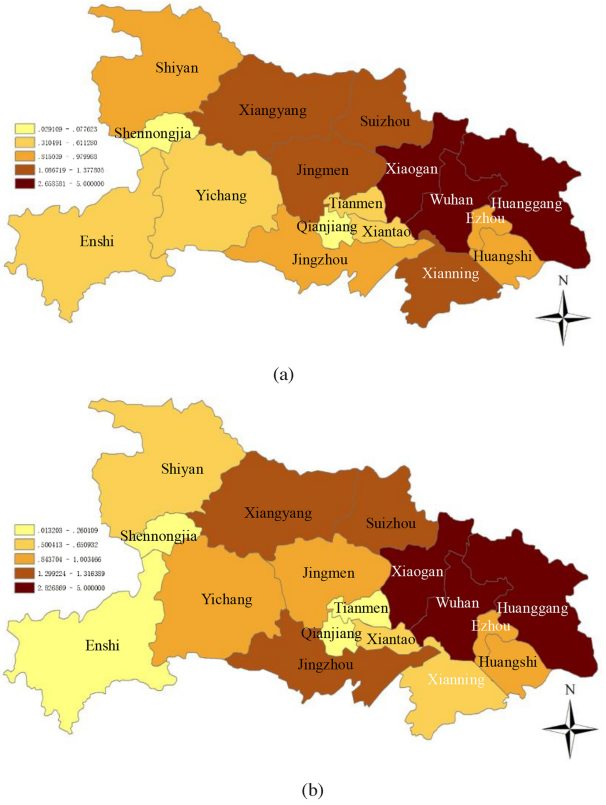

Fig. 8. Heat map of accumulated confirmed patients: (a) 28 January; (b) 8 February

Fig. 9. Heat map of total population inflow

(ii) most of the cities with higher prediction results are provincial-controlled divisions with smaller populations, while those with lower predicted values are mostly large cities. It is generally believed that cities with smaller populations have more advantages in travel control and epidemic prevention; (iii) almost all areas with low predicted values are located at the provincial boundary, and are therefore more susceptible to interprovincial migration. Furthermore, transportation hub cities may face a greater threat from an epidemic, due to their large traffic volume. For instance, population movements from east to west within the province pass through Jingmen, the central railway hub of Hubei, bringing a higher risk of human infection; and (iv) tourism was hit by the epidemic, with fewer tourists arriving than expected, although strict travel control was not adopted before 24 January. This accounted for the overestimated infections in the tourist cities, such as Enshi and Shennongjia.

5 Conclusions

In this paper, we calculated the population migration between cities using an improved gravity model. Using epidemic data and prediction results, a correlation between population migration and the epidemic spread of COVID-19 in major cities in Hubei Province was performed to show how migration is associated with epidemic

transmission. This research provides the following conclusions:

(i) Forces associated with intra-provincial human mobility included population, distance, socioeconomic development level, public services and resources, and traffic accessibility. Migration took place primarily from developed areas to underdeveloped areas before the Spring Festival, and the main destinations were the cities of Xiaogan, Huanggang, Ezhou and Huangshi.
(ii) Human migration had a significant impact on epidemic spread. Different functionary mechanisms existed in immigration from different regions: the mean correlation coefficient between the inflows from Wuhan and the number of infected people was 0.87, and that between the population inflow from cities except Wuhan and accumulated confirmed patients was 0.65. The total inflows accounted for epidemic transmission to a greater extent than Wuhan's immigration, despite the similarity in the correlation coefficients.
(iii) Compared with actual epidemic data, cities with small populations or tourist cities were overestimated, yet lower prediction results occurred in the interprovincial boundary region and traffic hub city. Additionally, the effect of distance on human mobility was amplified in the modified gravity model.

This study makes significant contributions to the current scholarly research. First, the existing literature has revealed that intra-provincial human mobility remains poorly understood. This study fills this important void by exploring which factors are more likely to dominate intra-provincial migration, with an eye toward the unique features of intra-provincial migration. Holding that public services and resources and traffic factors are critical to the choice of destination, a modified gravity model incorporating these was developed. Second, and relatedly, we highlight the time-dependent variations of migration with a direction coefficient, which increases the adaptability of the model. Urban status can be assessed by the migration networks associated with spatial pattern on population mobility. Third, many academics have proposed a close connection between the infection rate and human mobility. Echoing their argument, our findings demonstrate the direct impact of migration on epidemic transmission and deepen the understanding of the mechanism between these factors, suggesting the necessity of strict measures limiting travel nationwide, especially epidemic monitoring in the rural areas of cities around Wuhan. The geographical distribution of urban inflows and outflows is conducive to the analysis of the difference in epidemic spread between different cities. Furthermore, this research proves the feasibility of forecasting the development trend of COVID-19, in the case of missing epidemic data, which provides an essential reference for epidemic prevention and control.

Some limitations of this study should be noted. First, the empirical direction coefficient in a particular period and the distance-attenuation coefficient in the gravity model need to be verified by short-term migration data, which requires data collection and statistics. Another limitation lies with the epidemic data used for correlation analysis. Some cases of COVID-19 are not discovered and counted in time, given the existence of asymptomatic patients and the shortage of medical resources. Finally, the findings of this study may not reflect the impact of interprovincial migration, control policies and inner-city travel on disease transmission. Thus, we should be cautious in predicting the n numerical values of infections.

In view of these limitations, future research needs to gather information on intra-provincial human mobility with a significantly larger sample size to conduct the gravity-model coefficient calibration. Moreover, we have to collect adequate epidemic data to discuss the interaction between population movement and epidemic transmission with greater consideration for comprehensive influence mechanisms.

Nomenclature

M_{ij}	Intensity of population migration between city i and city j
Q_i	Quality of city i
Q_j	Quality of city j
A_{ij}	Comprehensive traffic accessibility between city i and city j
d_{ij}	Spatial distance between city i and city j
M'_{ij}	Population flow from city i to city j
M''_{ij}	Population flow from city j to city i
a	Direction coefficient of population migration
P_i	Number of permanent residents of city i
G_i	GDP per capita of city i
T_i	Proportion of tertiary industry in city i
h_{ij}	Accessibility of road network between city i and city j

r_{ij} Accessibility of rail network between city i and city j
a_{ij} Accessibility of civil aviation network between city i and city j
α Weight coefficients of roads in passenger-transport network
β Weight coefficients of railway in passenger-transport network
γ Weight coefficients of civil aviation in passenger-transport network
h'_{ij} Obstruction coefficient of road transportation between city i and city j
n^h_{ij} Number of roads connecting city i and city j
M'_i Outflows from city i
M''_i Inflows into city i
$\bar{d}_{ij}$ Result of normalized distance d_{ij}

Supplementary data

Supplementary data is available at *Transportation Safety and Environment* online.

Conflict of interest statement. None declared.

References

1. Sun K, Chen J, Viboud C. Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. *Lancet Digit Health* 2020; **2**:201–8.
2. Yang Z, Zeng Z, Wang K et al. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *J Thorac Dis* 2020; **12**:165–74.
3. Zhu N, Zhang D, Wang W et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 2020; **382**:727–33.
4. Zhang L, Liu S, Zhang G et al. Internal migration and the health of the returned population: a nationally representative study of China. *BMC Public Health* 2015; **15**:719.
5. Zipf G. The P1P2/D hypothesis: on the intercity movement of persons. *Am Sociol Rev* 1946; **11**:677–86.
6. Greenwood MJ Research on internal migration in the United States: a survey. *J Econ Lit* 1975; **13**:397–433.
7. Borjas GJ. Economic theory and international migration. *Int Migr Rev* 1989; **23**:457–85.
8. Lewis WA. Economic development with unlimited supply of labour. *Manch Sch* 1954; **22**:139–91.
9. Gupta MR. Migration, unemployment and development: a dynamic two-sector analysis. *Econ Lett* 1984; **16**: 177–84.
10. Richmond AH. Migration theory: talking across disciplines. *J Refug Stud* 2001; **3**:331–4.
11. DaVanzo J. Why families move: a model of the geographic mobility of married couples. *Popul Dev Rev* 1977; **3**: 344–5.
12. Hagen-Zanker J. Why do people migrate? A review of the theoretical literature. *Working Paper*. Maastricht Graduate School of Governance 2008.
13. Lee E. A theory of migration. *Demography* 1966; **3**:47–57.
14. Sjaastad L. The costs and returns of human migration. *J Polit Econ* 1962; **70**:80–93.
15. Crawford T. Beliefs about birth control: a consistency theory analysis. *Represent R Soc Psyc* 1973; **4**:53–65.
16. Taylor JE. The new economics of labour migration and the role of remittances in the migration process. *Int Migr* 1999; **37**:63–88.
17. Faist T. The crucial meso-level. In: Hammar T, Brochmann G, Tamas K et al. (eds). *International Migration, Immobility and Development: Multidisciplinary Perspectives*. Oxford: Berg, 1997, 187–217.
18. Clark GL, Ballard KP. The demand and supply of labor and interstate relative wages: an empirical analysis. *Econ Geogr* 1981; **57**:95–112.
19. Bluestone B, Harrison B. The deindustrialization of America. *Bus Horiz* 1983; **26**:80–1.
20. Liang Z, White MJ. Market transition, government policies, and interprovincial migration in China. *Econ Dev Cult Change* 1997; **45**:1983–8.
21. Duan CR. Individual level determinants of inter-provincial migration in China: on the effects of time sequence in migration studies. *Popul Res* 2000; **24**:14–22.
22. Tang JL, Ma ZD. Migration selectivity in China: an analysis based on the 2000 census data. *Popul Res* 2007; **31**: 42–51.
23. Yuan X, Wang H, Chen Y. Analysis on factors affecting decision-making of migrant population in large cities of China. *Population Journal* 2011; **3**:59–66.
24. Fan CC. Interprovincial migration, population redistribution, and regional development in China: 1990 and 2000 census comparison. *Prof Geogr* 2005;**57**:295–311.
25. González MC, Hidalgo CA, Barabási A-L. Understanding individual human mobility patterns. *Nature* 2008; **453**: 779–82.
26. Zhang Y, Cen Q. Spatial patterns of population mobility and determinants of inter-provincial migration in China. *Popul Res* 2014; **38**:54–71.
27. Jiang X, Wang S. Research on China's urban population mobility network: based on Baidu migration big data. *Chin J Popul Sci* 2017; **1**:35–46.
28. Mansell PR. Disease and mobility: a neglected factor in epidemiology. *Int J Epidemiol* 1977; **3**:259–67.
29. Martens P, Hall L. Malaria on the move: human population movement and malaria transmission. *Emerg Infect Dis* 2000; **6**:103–9.
30. Neouimine NI Leishmaniasis in the eastern Mediterranean region. *East Mediterr Health J* 1996; **25**:66–8.
31. Gushulak BD, MacPherson DW. Population mobility and infectious diseases: the diminishing impact of classical infectious diseases and new approaches for the 21st century. *Clin Infect Dis* 2000; **31**:776–80.
32. Garnett GP, Lewis JJC. The impact of population growth on the epidemiology and evolution of infectious diseases. In: Caraël M, Glynn JR (eds). *HIV, Resurgent Infections and Population Change in Africa*. Dordrecht: Springer, 2007, 27–40.
33. Gani JM, Yakowitz SJ, Blount M. The spread and quarantine of HIV infection in a prison system. *SIAM J Appl Math* 1997; **57**:1510–30.
34. Brauer F, Driessche PVD. Models for transmission of disease with immigration of infectives. *Math Biosci* 2001; **171**: 143–54.
35. Wang W, Mulone G. Threshold of disease transmission in a patch environment. *J Math Anal Appl* 2003; **285**:321–35.

36. Takeuchi Y, Liu X, Cui J. Global dynamics of SIS models with transport-related infection. *J Math Anal Appl* 2007; **329**:1460–71.
37. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modeling study. *Obstet Gynecol Surv* 2020; **75**:399–400.
38. Tang B, Wang X, Li Q et al. Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. *J Clin Med* 2020; **9**:462–74.
39. Read JM, Bridgen JRE, Cummings DAT et al. Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic prediction. 2020. medRxiv, 10.1101/2019.12.11.12345678.
40. Tang SY, Tang B, Bragazzi Nl et al. Analysis of COVID-19 epidemic traced data and stochastic discrete transmission dynamic model. *Sci China Math* 2020; **50**:1071–86.
41. Xu XK, Wen C, Zhong GY et al. The geographical destination distribution and effect of outflow population of Wuhan when the outbreak of the COVID-19. *J Univ Electr Sci Technol China* 2020; **1**:1–6.
42. Liu Y, Yang DY, Dong GP et al. The spatio-temporal spread characteristics of 2019 novel coronavirus pneumonia and risk assessment based on population movement in Henan Province: analysis of 1243 individual case reports. *Econ Geogr* 2020; **3**:24–32.
43. Liu T, Qi YJ, Cao GZ et al. Spatial patterns, driving forces, and urbanization effects of China's internal migration: county-level analysis based on the 2000 and 2010 censuses. *J Geog Sci* 2015; **25**:236–56.
44. Wilson ME. Travel and the emergence of infectious diseases. *Emerg Infect Dis* 1995;**1**:39–42.
45. Tian PP, Zhu Y, Lin LY et al. Differences in the spatial distribution and its determinants between inter- and intra-provincial floating population: the case of Fujian Province. *Population Journal* 2015; **37**:56–67.
46. Yan XY, Wang WX, Gao ZY et al. Universal model of individual and population mobility on diverse spatial scales. *Nat Commun* 2017; **8**:1639–48.
47. Li Q, Guan X, Wu P et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med* 2020; **382**:1199–207.
48. Ravenstein EG. The laws of migration. *J Stat Soc London* 1885; **48**:167–235.
49. Greenwood MJEA. Migration, regional equilibrium, and the estimation of compensating differentials. *Am Econ Rev* 1991; **81**:1382–90.
50. Stratford D. Estimating relative standard of living in the United States using cross-migration data. *J Regional Sci* 1997; **37**:411–36.
51. Shen J. Modelling regional migration in China: estimation and decomposition. *Environ Plan A* 1999; **31**: 1223–38.
52. Duan CR, Lyu LD, Wang H et al. From rural China to migrating China: rethinking migration transition in China. *Popul Res* 2020; **44**:19–25.
53. Bao SM, Shi AN, Hou WZ. Analysis of the changing spatial patterns of migration in China. *Popul Sci China* 2005; **95**:56–61.
54. Peeters D, Thomas I. Distance predicting functions and applied location-allocation models. *J Geogr Syst* 2000; **2**: 167–84.
55. Chi G. The impacts of highway expansion on population change: an integrated spatial approach. *Rural Sociol* 2010; **75**:58–89.
56. Li T, Cao XS, Huang XY. The relationship between spatial structure of accessibility and population change in Pearl River Delta. *Geogr Res* 2012; **9**:1661–72.
57. Zhao ZY, Wei Z, Yang R et al. Gravity model coefficient calibration and error estimation: based on Chinese interprovincial population flow. *Acta Geogr Sinica* 2019; **74**: 203–21.
58. Yan SP. Research on the mechanism of China's interprovincial population migration. *Chin J Popul Sci* 2007; **1**: 71–77.
59. Guan WJ, Ni ZY, Hu Y et al. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020; **382**: 1708–20.