ARTICLE

Check for updates

Examining the population flow network in China and its implications for epidemic control based on Baidu migration data

Sheng Wei[1] & Lei Wang [2,3✉]

This paper examines the spatial pattern of the population flow network and its implications for containing epidemic spread in China. The hierarchical and spatial subnetwork structure of national population movement networks is analysed by using Baidu migration data before and during the Chinese Spring Festival. The results show that the population flow was mainly concentrated on the east side of the Hu Huanyong Line, a national east-west division of population density. Some local hot spots of migration were formed in various regions. Although there were a large number of migrants in eastern regions, they tended to concentrate in corresponding provincial capital cities and the population movement subnetworks were affected by provincial administrative divisions. The patterns identified are helpful for the provincial government to formulate population policies on epidemic control. The movement flow from Wuhan (the city where the covid-19 outbreak) to other cities is significantly and positively correlated with the number of confirmed cases in other Chinese cities (about 70% of the population was constituted through innerprovincial movement in Hubei). The results show that the population flow network has great significance for informing the containment of the epidemic spread in the early stage. It suggests the importance for the Chinese government to implement provincial and municipal lockdown measures to contain the epidemic spread. The paper indicates that spatial analysis of population flow network has practical implications for controlling epidemic outbreaks.

[1] Jiangsu Institute of Urban Planning and Design, Nanjing, China. [2] Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China. [3] Planning and Environmental Management, School of Environment, Education and Development, University of Manchester, Manchester, UK. ✉email: wanglei@niglas.ac.cn

Introduction

The spatial pattern of population movements reflects the relationship between cities and constitutes important information in exploring the economic connections and traffic demands between cities (Shumway and Otterstrom, 2010; Gonzàlez et al., 2008; Ravenstein, 1884). With the diversification and efficient development of different modes of transportation, population flow patterns are increasingly examined in terms of complexity and diversity (Wei et al., 2019). For a country with a large land area like China, analysing the pattern of population flow during a specific period is of great significance for understanding economic and social development, in addition to the control of epidemic outbreaks (Read et al., 2020; Bajardi et al., 2011; Colizza et al., 2006).

Examining spatial patterns of population flow requires the deployment of new methods for data collection in ongoing research (Shen, 2020). Many recent studies have researched population movement based on mobile phone data, which has the benefit of high accuracy and large sample size (Bengtsson et al., 2015). These studies have explored geographic borders of human mobility, individual human mobility patterns, travel behaviour, and the spatial structure of cities (Gariazzo and Pelliccioni, 2019; Lee et al., 2018; Picornell et al., 2015; Sagl et al., 2014; Louail et al., 2014; Rinzivillo et al., 2012). As mobile phone data are generally expensive and there are privacy concerns, researchers encounter many obstacles in the collection of large amounts of data for use in large-scale spatial population flow analysis (Wei et al., 2017). Fortunately, open big data sources provide new research opportunities and increase the credibility of the results. For example, the frequency data of trains (Wang, 2018) and flights (Guimera et al., 2005) were used to understand the trends in population movement between cities. Geo-tagged social media data have also been used to explore the characteristics of human mobility at a much finer temporary and spatial scale (Khan et al., 2020; Hawelka et al., 2014; Naaman et al., 2012; Kamath et al., 2012).

Generally, research on Chinese population flow networks encounters several problems with respect to data sources. Acquiring national-scale data is difficult and the data sources are not uniform. In this study, migration data from the Baidu map service, a Chinese equivalent to Google, were used to study the spatial pattern of population flow. Such open big data have three apparent advantages when compared with traditional data sources. It contains national population flow data, which allows for a dynamic study of large-scale population flow (Deville et al., 2014). It also provides a relatively uniform type of data as it comes from the external release results of a single Chinese internet company. The data is generated through mobile phone software with many users, and thus, the results of the data analysis are more reliable (Pei et al., 2014).

While studying the characteristics of population flow between cities, we often treat cities as network nodes and use the population flow between two cities as the edge weight of the network node. Therefore, the complex network theory (Watts and Strongatz, 1998; Barabasi and Albert, 1999) can be used to study such data, because it accurately represents the relationships between network nodes and evaluating network node results from different perspectives (Ding et al., 2019). The spatial pattern of population flow networks could be applied to examine epidemic transmission as it involves the lifestyles and physical health of migrating people.

Each year during the Spring Festival, there is a large-scale national population movement in China as people return to their hometowns for New Year celebrations and family reunions. Travel during the Spring Festival is called 'Chunyun', which is an annual exam of the Chinese transportation systems. In 2020, Chunyun officially began on 10 January. As of 12 February, owing

to population movement, covid-19, which outbreak in Wuhan, Hubei Province, China at the end of 2019, spread throughout China (Shen, 2020; Chung et al., 2020). On 23 January 2020, Wuhan implemented a lockdown policy to cut down on the spread of the disease. Therefore, the Baidu migration data between 10 and 22 January 2020 reflects the impact of Chunyun on the transmission and spread of covid-19. Using Baidu migration data from 1 to 9 January, we can obtain the general spatial patterns of population movements in China before Chunyun (the non-Chunyun period). Consequently, this paper aims to understand how covid-19 spread across China by using the population flow network as a predictor. The research questions were examined through the following three interrelated perspectives:

First, we examined the hierarchical structure of population flow to understand the role of important node cities in the non-Chunyun period. In the literature, Liu and Gu (2019) explored the spatial concentration patterns of Chinese interprovincial population migration using the interprovincial population between origins and destinations (the OD indicator). The hierarchical structure of population flow network nodes was assessed using two network assessment indicators, namely weighted degree and betweenness centrality (Wang et al., 2020; Wei et al., 2019; Shanmukhappa et al., 2018). In this study, we also used the two indicators to examine the important node cities in the population flow network. Second, we identified subnetworks of Chinese population flow and their spatial distribution characteristics in the non-Chunyun period as a baseline to refer to the spread of covid-19. In recent years, the Chinese government has been promoting urbanisation in metropolitan areas as its core strategy (Yin et al., 2017). Urban agglomeration is a crucial area to attract migrants. Based on the Baidu migration data from January 1 to January 9, 2020, we explored the closely related subnetworks through the community detection method, thus providing a spatial reference for identifying urban agglomerations and its role in covid-19 expansion. Third, the case of covid-19 was examined to explore the role of population flow in the spread of the epidemic during Chunyun. By analysing the correlation between the migration volumes from Wuhan and the number of covid-19 cases in other cities, we explored the impact of population movement on infectious diseases. Besides, the relationship between the number of migrants from Wuhan and the covid-19 cases in other cities in Hubei Province was analysed through the approach of Chord Diagram. The Chord Diagram comprises nodes, arc lengths corresponding to nodes, and chords among the nodes (Gu et al., 2014), and it can quickly reveal the relationship and patterns among the origins, destinations, flow directions, and connection intensities of the population flow (Liu and Gu, 2019; Qi et al., 2017). These results are useful for population management and emergency response during certain periods (Balcan et al., 2009).

The rest of this paper is arranged as follows. The second section provides an introduction to the methodology and data sources relied on and is followed by the results of the spatial pattern analysis of the Chinese population flow network in section three. The last section concludes the study.

Research materials and methods

Data pre-processing and network construction. The Baidu migration data were acquired from Baidu Map (http://qianxi.baidu.com/), which is one of the most popular map service providers in China. It provides real-time location searching services through the GPS. Data pre-processing proceeded as follows. First, after data cleaning, we collected the migration scale index
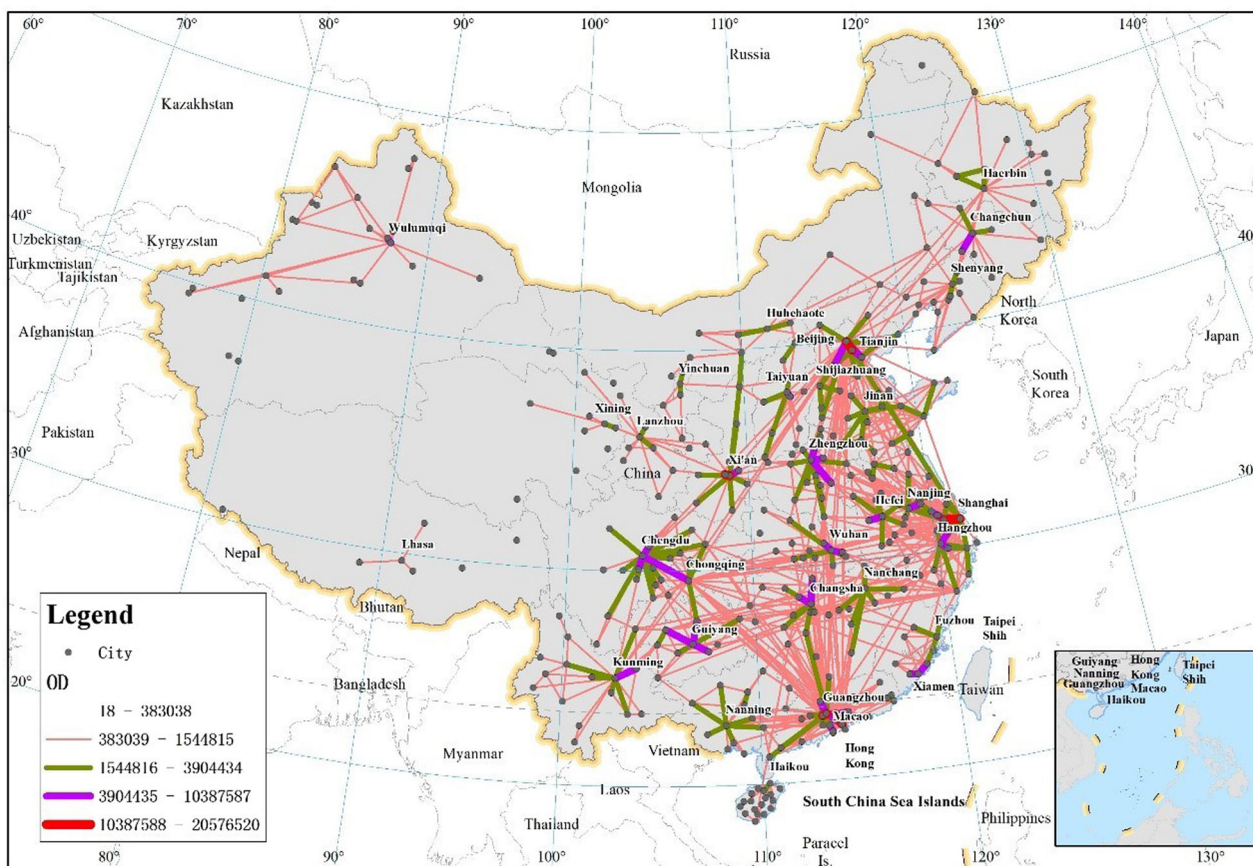
Fig. 1 The spatial distribution of OD during the non-Chunyun period. The width of lines represents the migration volume from cities to cities.

data between any two cities among 367 major cities in China. The migration scale index data record the number of migrants between any two cities. It visualises the migration flows in particular time periods. Although it cannot capture all migrants as it is based on the availability of smartphones, it is useful to conduct the comparisons among different cities in general migration patterns. The migration scale index data between two cities is used to indicate the intensity of population movement, and this index has an origin-destination (OD) record from one city to another. Second, based on the complex network theory and its approach, each city is considered a network node. Once there is an OD record between any two cities, those two cities are considered to have an edge connection. The record value is considered the edge weight between those two nodes. Then, we multiplied each record value by 100 to convert all records into integers. This generated 433,787 OD records between 367 cities from 1 to 22 January 2020. The source of the data is divided into two parts. Part one extends from 1 to 9 January, and was used to analyse the characteristics of population flow pattern during the non-Chunyun period. Part two extends from 10 to 22 January and is used to analyse the population flow pattern during the Chunyun period.

The number of confirmed covid-19 cases in different cities came from the Tianditu map website (https://www.tianditu.gov.cn/coronavirusmap/) released by China's Centres for Disease Control. Considering that the incubation period of covid-19 is about 14 days, we obtained the number of confirmed patients as of 6 February 2020 (14 days after the lockdown policy) for each city.

Methodology

Weighted degree centrality. The degree centrality of a node in a complex network is defined as the number of edges associated with a given network node. Therefore, the degree centrality describes the state of the connection among the nodes in a network. The formula for calculating this network index is as follows:

$$D_i = \sum_{j=1, j \neq i}^{n} L_{ij} \tag{1}$$

where, L_{ij} is the number of direct connections between i and j, and n is the total number of nodes.

Weighted degree centrality (WDC) is the degree obtained once the weight between any two network nodes in an urban bus network has been considered (Newman, 2001). Generally, the WDC can directly reflect the hierarchical status of nodes in the population network. The formula for calculating the WDC is as follows:

$$S_i = \sum_{j \in N_i} w_{ij} \tag{2}$$

where N_i is the set of nodes adjacent to node I, and W_{ij} is the weight of the edge between nodes i and j.

Betweenness centrality. Betweenness centrality (BC) is defined as the number of shortest paths between two network nodes that pass through the current network node. It is used to measure the importance of nodes acting as bridges in a network (Freeman, 1977, 1979; Newman, 2001) and is mathematically defined by
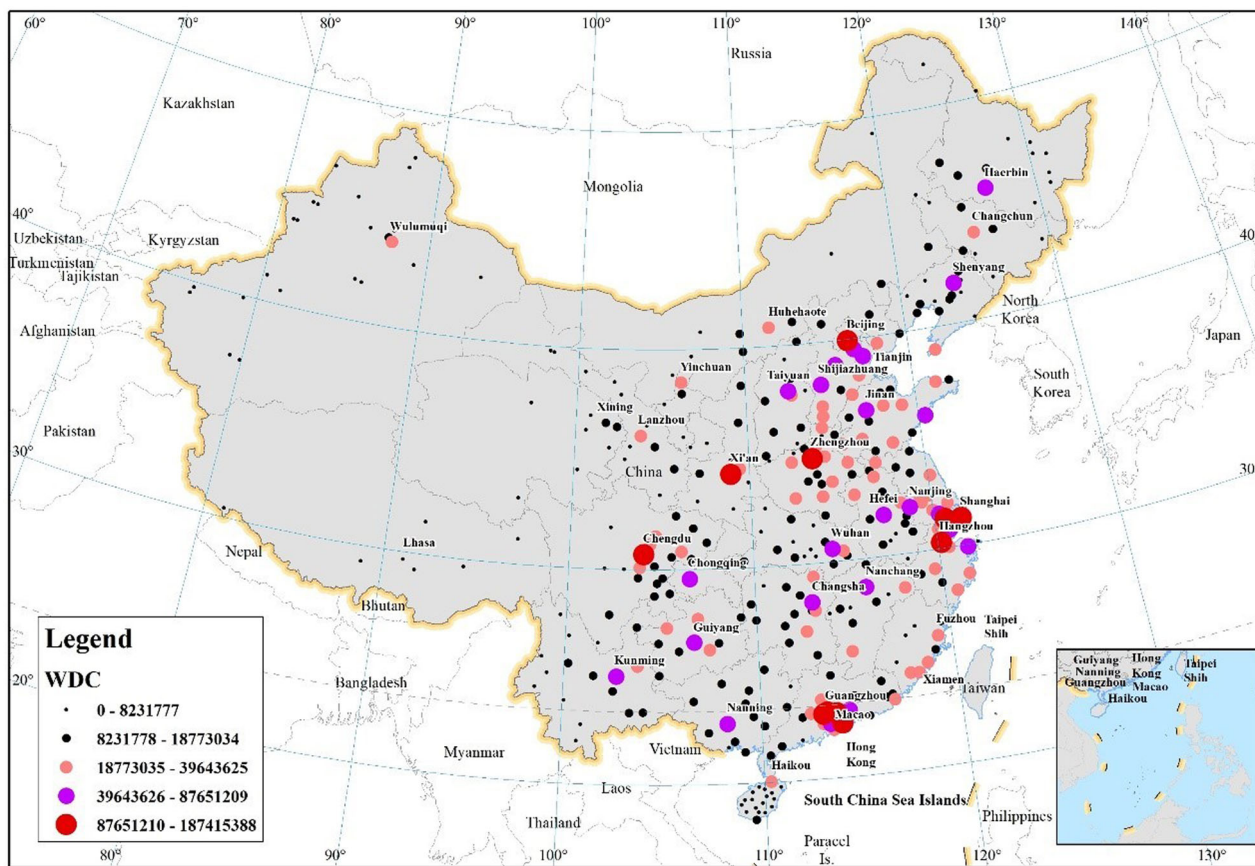
Fig. 2 The spatial distribution of WDC during the non-Chunyun period. The size of bubbles represents cities' weighted degree centrality in the national population flow network.

Eq. (3), where s and t comprise a pair of clustered nodes, V, $\sigma(s, t)$ indicates the number of the shortest paths between nodes s and t, and $(s, t|i)$ is the number of the shortest paths between nodes s and t passing through node i.

$$B_i = \sum_{s,t \in V;\, s,t \neq i} \frac{\sigma(s, t \mid i)}{\sigma(s, t)} \tag{3}$$

Subnetwork detection. Subnetwork detection is a method used to divide an entire network into closely connected subnetworks, which can be understood as a group of network nodes with some shared characteristics. Modularity is a measure of the level or degree to which a network's subnetworks may be separated and recombined, which is a commonly used criterion for determining the quality of network partitions. The larger the modular value, the better the division of the subnetwork structure. The upper limit of the modular value is 1, wherein the closer the value is to 1, the better the subnetwork structure. In a real network, the value of modularity usually ranges from 0.3 to 0.7 (Chen et al., 2012).

Results
The hierarchical structure of population flows. Figure 1 shows the uneven global distribution of population flow in China during the non Chunyun period. It is mainly concentrated on the east side of the Hu Huanyong Line, a well-recognised east-west divide of population density in China (Hu, 1935). The line has been repeatedly identified as a climate-driven separation of densely and sparsely populated areas in China (Hu et al., 2016). With one or more provincial capitals as the centre, some areas of concentrated population flow have been formed. These areas were scattered across various regions. Except for the economic and strategic

importance of provincial capital cities (such as Chengdu, Chongqing, Shanghai, and Hangzhou), there is also a high intensity of floating populations between the provincial capital cities and regional central municipalities.

As shown in Fig. 2, the spatial distribution of WDC reveals the hierarchical structure in the Chinese population flow network, and a series of core cities has been formed in the population flow network. Among these core cities, Chengdu, Zhengzhou and Xi'an which are located in central and western China were quite attractive as destinations for prospective migrants. However, they are not in the three most developed urban agglomerations of Yangtze River Delta, Beijing-Tianjin-Hebei, and Pearl River Delta in China (Liu et al., 2016; Liu and Wang, 2016). The pattern was partly related to the industrial transfer from Eastern to Central China in recent years (Wang, et al., 2020).

As shown in Fig. 3, the spatial distribution of BC reflects the distribution of cities that has an important ability to control the population flow network. This is also understood as a hub in the urban network. We find that the spatial distribution characteristics of BC are associated with those of WDC in general. Among them, Wuhan is ranked 6th in terms of both the WDC and BC indicators, suggesting that Wuhan plays an essential role in the population flow network in the country. In other words, if the covid-19 outbreak had occurred during the non-Chunyun period, Wuhan might still have had a significant role to play in the spread of the disease.

The scatter diagram between WDC and BC is shown in Fig. 4. Cities with larger population flow also have a stronger capacity to control the population flow network. Some cities with large floating populations have different characteristics in the relationship between WDC and BC. Compared with the cities in area A
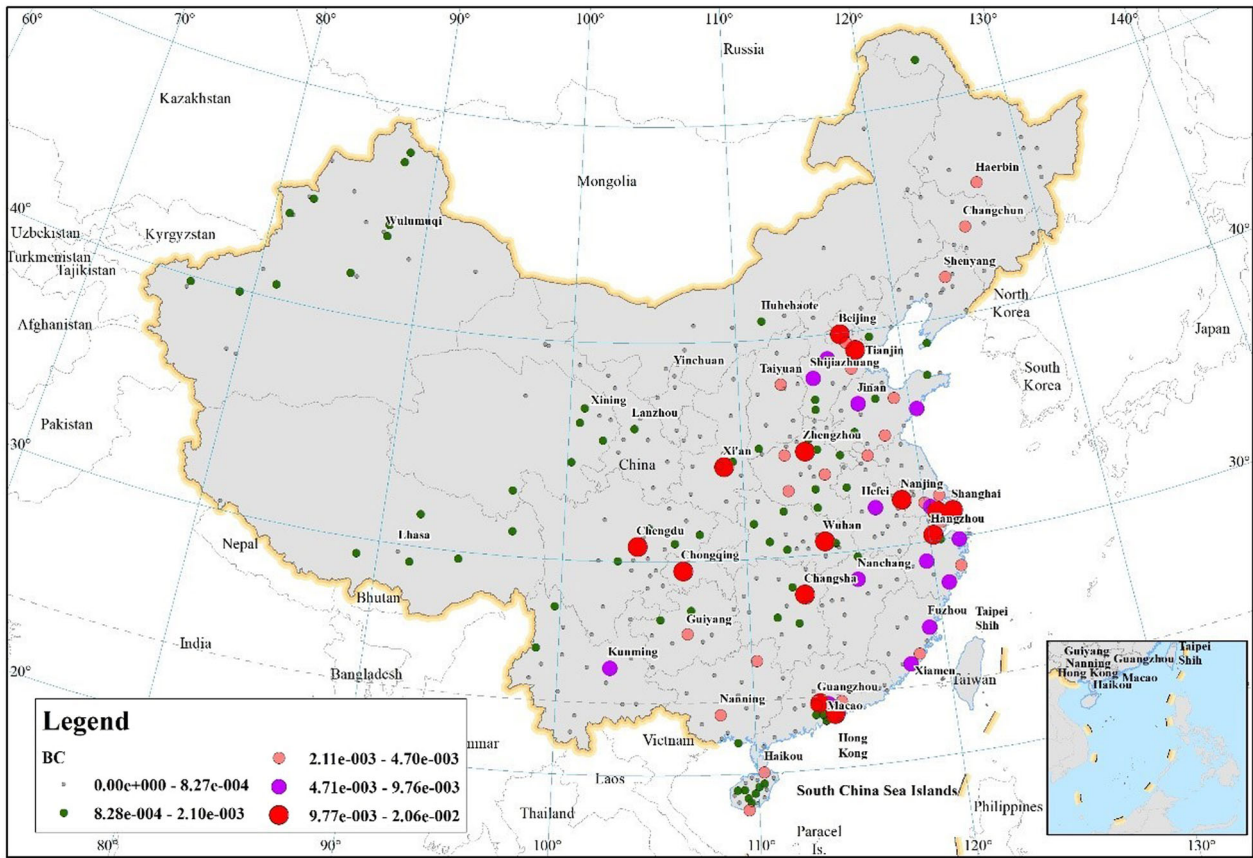
Fig. 3 The spatial distribution of BC during the non-Chunyun period. The size of bubbles represents cities' betweenness centrality in the national population flow network.
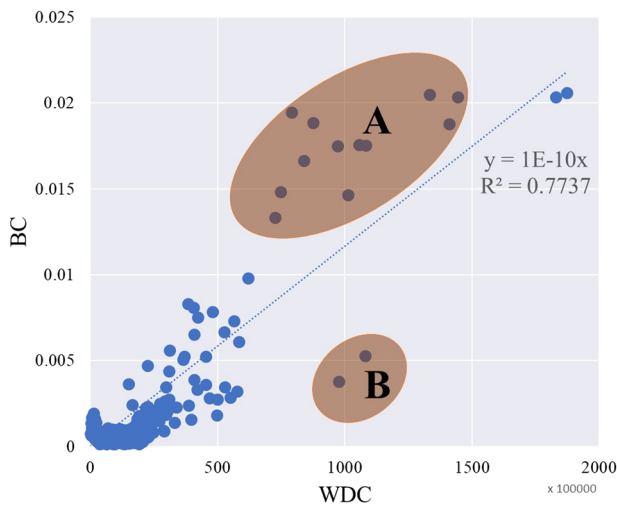

Fig. 4 The scatter diagram between Ln(WDC) and BC during the non-Chunyun period. Scatter plot of WDC and BC.

in Fig. 4, the two cities (Foshan and Dongguan) in area B have a low BC value despite having large floating populations. Here, we take Dongguan as an example to explain the reason for this phenomenon. Dongguan is close to Shenzhen, an economically developed Chinese city. In recent years, owing to the shortage of land resources and rising housing prices in Shenzhen, many enterprises have chosen to set up factories or branches in Dongguan. This development process requires high-quality transport services to make frequent roundtrips between Dongguan

and Shenzhen. Thus, a typical feature of these two cities is that they had a large portion of the total migrations with a few cities.

The subnetwork of population flows. As shown in Fig. 5, the population flow network is divided into 12 subnetworks by the community detection method. The boundaries of those subnetworks roughly coincide with the provincial boundaries. This phenomenon demonstrates a close relationship between the population movements and administrative divisions. Although most subnetworks comprise cities of the entire province, while a few comprise multiple ones and a part of cities in the province. Overall, geographic proximity is a significant predictor of subnetwork composition. The spatial distribution characteristics of migration subnetworks could be used to inform the Chinese government on relevant population policies.

The spatial distribution of population movement OD between any two cities, where each of them belongs to a different subnetwork, is portrayed in Fig. 6. It shows that the primary structure was formed by the metropolitans between Beijing, Shanghai, Guangzhou, Chengdu-Chongqing, and their neighbouring cities. This main pattern had a strong connection with four periphery cities (Changchun, Harbin, Shenyang, and Hohhot) in northern China. Within the primary pattern, there were high-intensity population flows. The identified population movement patterns reflect the need of transportation services among regions and cities. Due to the long distances, the traffic links between some non-adjacent subnetworks mainly depend on the aviation network.

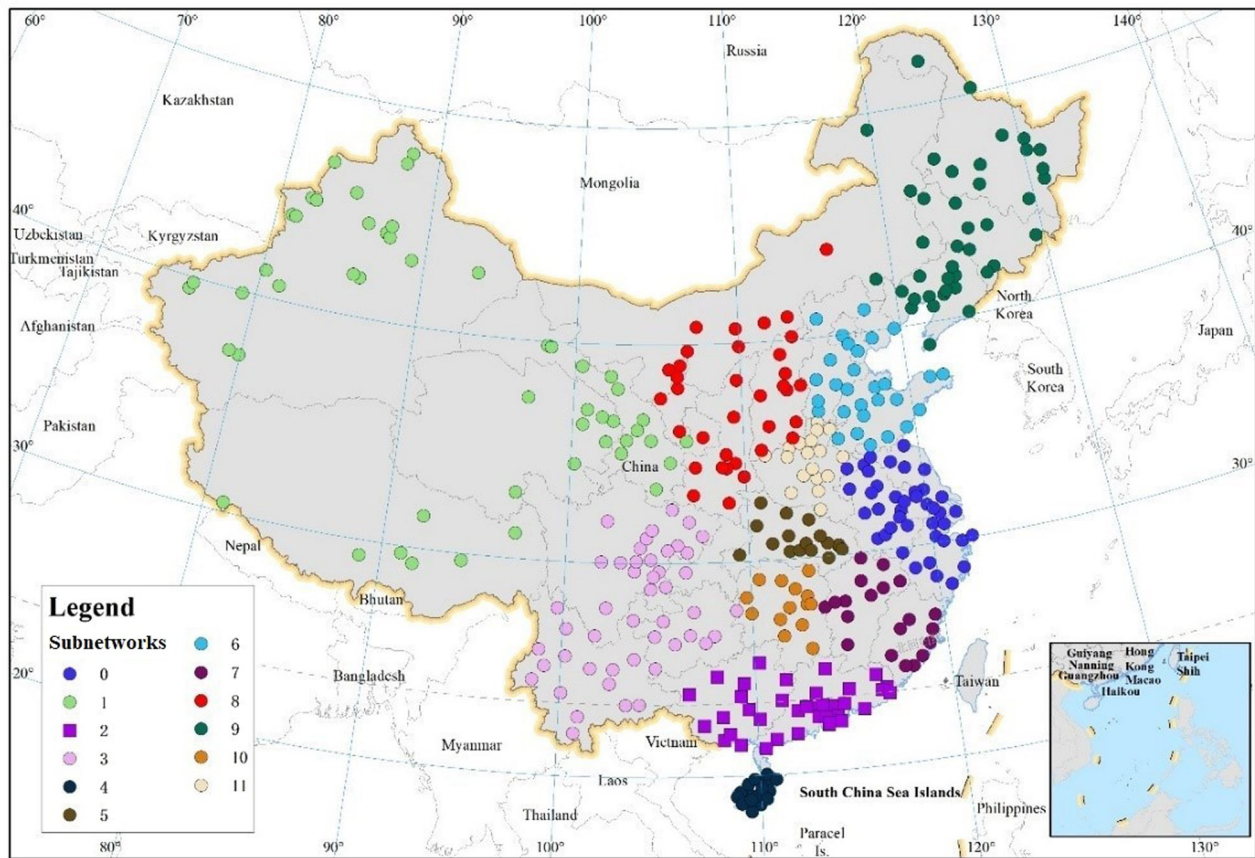Meanwhile, China is rapidly building a nationwide high-speed rail network system to meet the needs of rapid population

Fig. 5 The spatial distribution of subnetworks during the non-Chunyun period. The color of bubbles represents cities belonging to different subnetwork in the national population flow network.

movement. Therefore, aviation transport networks are facing tremendous competitive pressure from high-speed railway networks. As the two modes have their own optimal distances to attract sufficient passengers, our results can serve as a reference for planning population movement and transport networks.

The case of the spread of covid-19. We first examine the channel and pattern of the spread of covid-19 based on the population movement network before the Chunyun period. Figure 7 shows that Hubei Province, with Wuhan as its capital, comprises a single population flow subnetwork. Thus, compared with cities outside Hubei Province, Wuhan has a far greater population flow with other cities within the province. From the perspective of the impact of population flow on pandemic transmission, the prevention and control pressure in Hubei Province is more significant than in other provinces.

Second, we explore the relationship between the migration volumes and the spread of the epidemic between Wuhan and other cities in Hubei Province. Among the 17 cities and regions administrated directly by the provincial government, Shennongjia in Hubei Province is a protected forest area with only a population of 80,000 residents, so the impact of the epidemic on the area is relatively small. Figure 8a shows the ranking of the number of migrants from Wuhan to other cities through a Chord Diagram. We find that the top 17 cities, except Chongqing and Xinyang, were all cities in Hubei Province. Figure 8b also reveals that the migrating population from Wuhan flows mainly to cities in Hubei Province, accounting for 68.93%. Thus, the population flow in Wuhan is mainly concentrated in cities within the

province, particularly in Wuhan's neighbouring cities of Xiaogan and Huanggang. Figure 9 reveals that spatial distance is a significant factor affecting the spread of covid-19. The epidemic worsened in cities close to Wuhan. Correspondingly, the Chinese government implemented a lockdown policy in Wuhan and Huanggang and put stringent early controls in place on the population flow between Hubei and other provinces to mitigate the spread of the epidemic nationwide.

Third, as shown in Fig. 10, the number of migrants from Wuhan is significantly and positively correlated with the number of confirmed covid-19 cases in other cities. Wuhan is a traditional national education centre with several key national universities. It ranks among the top 5 cities with the number of well-known universities in China. The city accommodates nearly 1 million college students and ranked second among Chinese cities in this context in 2018. Migrant workers and college students return home for their holidays during the Spring Festival, and this might have increased the spread of covid-19.

Some details about the movement features could be explored by the example of Wenzhou. Although the floating population from Wuhan to Wenzhou was relatively low, the number of covid-19 cases diagnosed was relatively high (Fig. 10). Many people in Wenzhou are engaged in business and other service industries across the country. The population in Wenzhou travels for business more than those in other areas. People from business sectors are more likely to interact with others often (Liu and Xu, 2017; Liu et al., 2016), which may have increased their chances of contracting covid-19. This accounts for the disparity of Wenzhou in the spread of covid-19. It suggests that business travel and
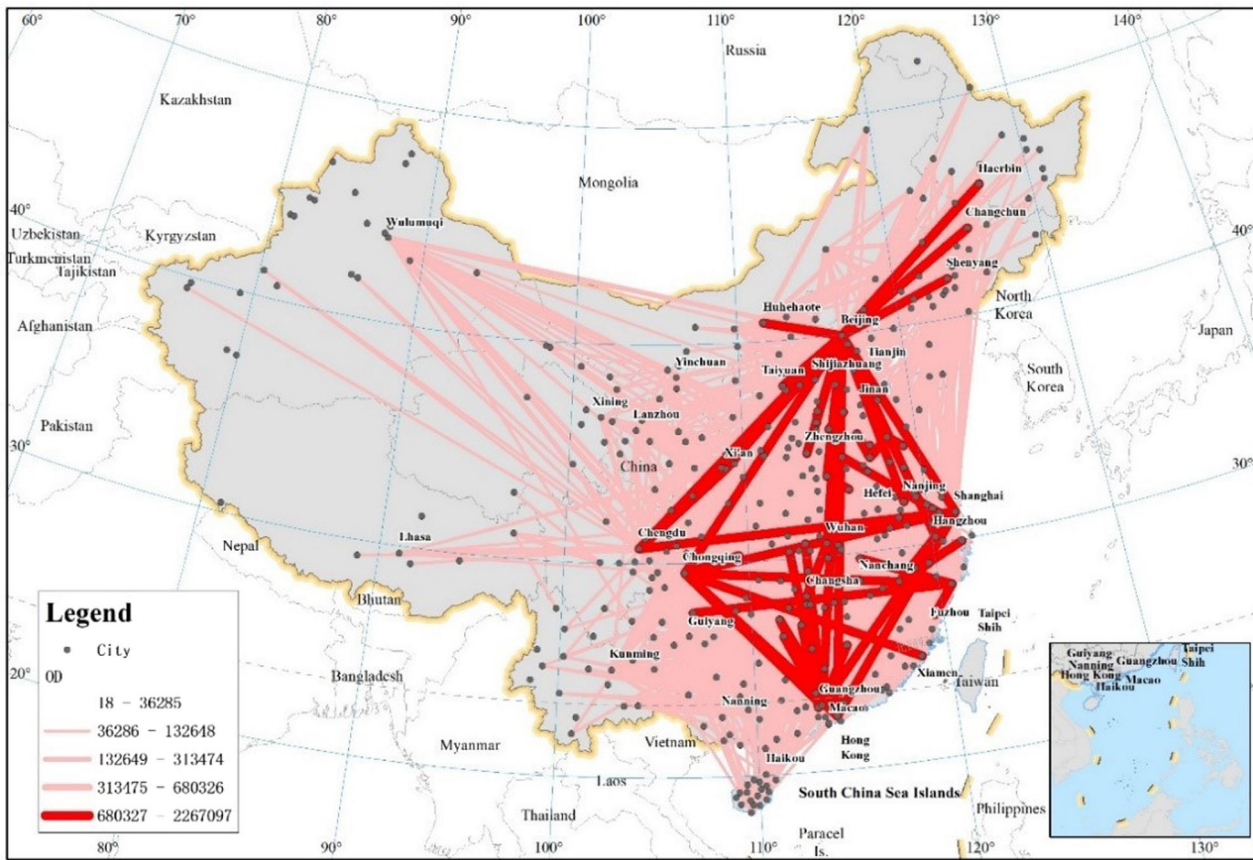
Fig. 6 The spatial distribution of OD between any two cities, and each belongs to different subnetworks during the non-Chunyun period. The width of lines represents the migration volume from cities to cities.

migration during the Chunyun period played important roles in the spread of covid-19 in China.

Discussions and conclusions

Different from the census data, the Baidu migration data are the most updated source for analysing the spatial pattern of population flow, and the findings are relevant to evidence-based regional population policymaking in highly mobile societies (Ma et al., 2015). This paper examined the pattern of Chinese population flow before and during Chunyun and its implications for the containment of covid-19 expansion. The spatial distribution of the population flow network was examined to confirm the critical role of provincial capitals in the Chinese migration pattern. The results suggest that multiple regional core cities have been formed, and this is conducive to the balanced development of a country like China with a vast territory. Based on the spatial distribution of WDC and BC, we obtained the hierarchical structures of Chinese cities from two different perspectives. Overall, cities with higher population flow also have stronger control in the population flow network. However, cities with a superior geographical location can help improve their dominance in the population flow network. Although some cities have a relatively large number of migrants, they had a relatively lower hub position in the population flow network because the number of migrants mainly comes from a small number of cities.

The community detection result reveals the distribution and restructuring of subnetworks in the population flow network before and during the Chunyun periods. Furthermore, from the spatial distribution of population flow between cities (each belongs to a different subnetwork), we revealed the characteristics

of major interprovincial population movements. Finally, we found that the migrating population of Wuhan was mainly related to cities in Hubei Province through the community detection method and a Chord Diagram. The statistical analysis showed that the intensity of the population flow relationship between cities would affect the spread of the transmission of a disease to a large extent.

The analysis in this paper has some spatial implications on population policy and epidemic control. For example, the spatial analysis of population movement is important for evidence-based population policymaking. The population flow is a pertinent indicator for the evaluation of the future development prospects of a city and is of great significance for the prediction and prevention of the transmission of a disease. These findings illustrate the importance of building an emergency system based on big data technology (Brooks et al., 2008).

Some limitations remain to be addressed in future research. For example, the longitudinal population flow data should be collected and used to capture the holistic migration pattern. The relationship between population flow and different transportation modes and the impact of population movement on industrial development can be examined to guide regional planning and development. Mobile signalling data (Calabrese et al., 2010; Widhalm et al., 2015) can be used to trace the spreading path of epidemic transmission at the city scale. The Baidu migration data should also be validated as some travellers, such as children and seniors, may not have smartphones and may not be included in the sample. The data on train and/or bus ticket sales can also be used to triangulate the data to ascertain reliability. Future research can use similar methods and sources of data to explore these limitations and draw out other practical benefits for
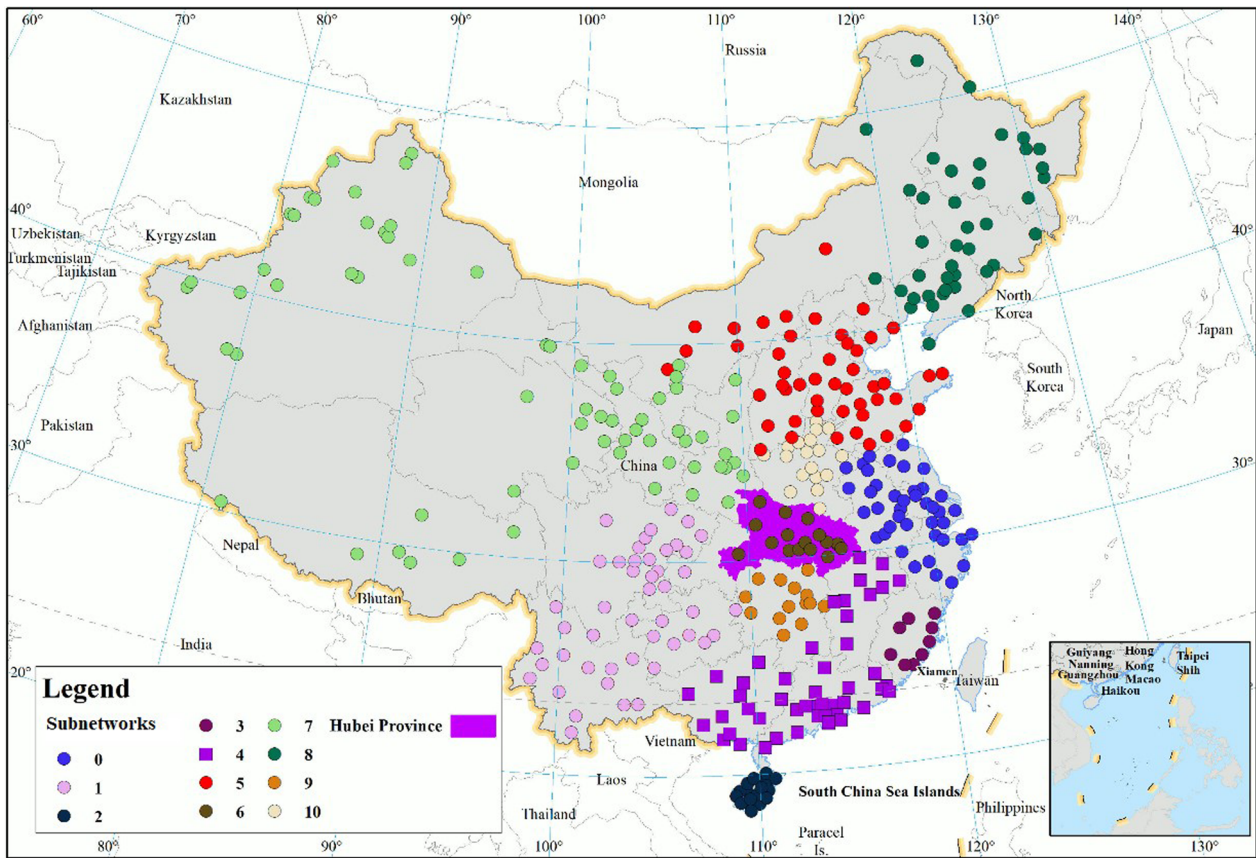
Fig. 7 The spatial distribution of subnetworks during the Chunyun period. The color of bubbles represents cities belonging to different subnetwork in the national population flow network. Hubei province is highlighted in purple.
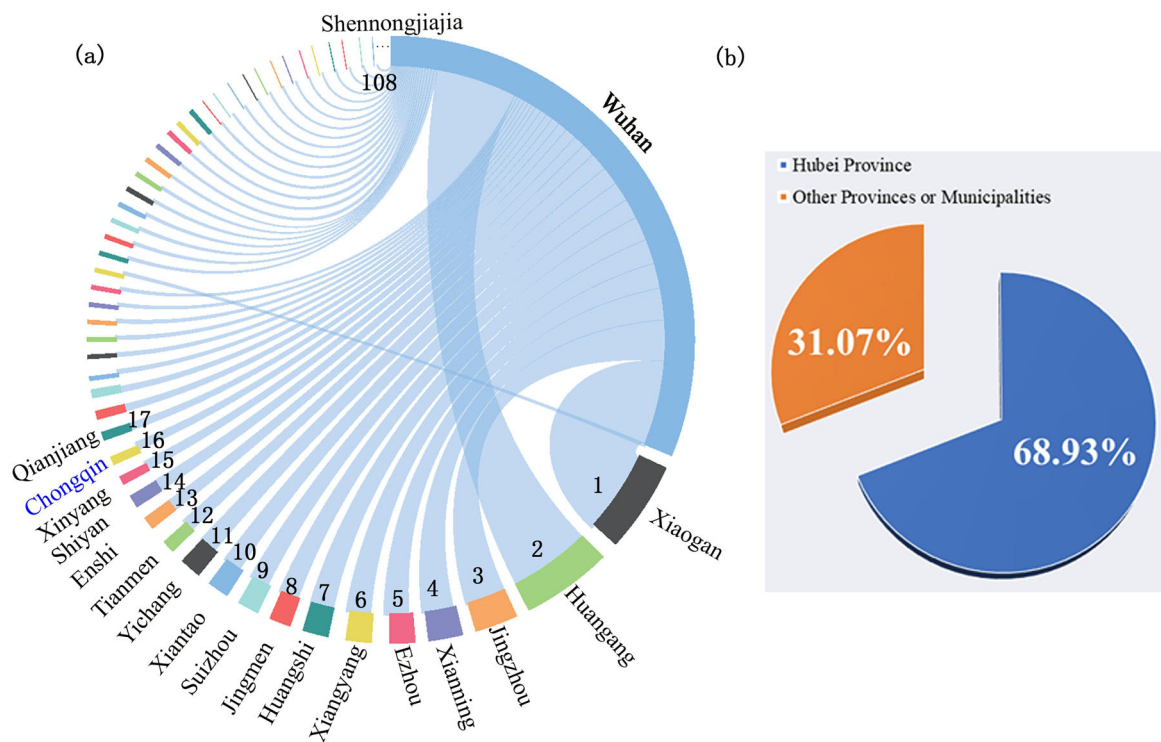

Fig. 8 a The number of migrants from Wuhan to other cities (TOP 50) during Chunyun. **b** Comparison between the number of migrants from Wuhan to other cities within and outside the Hubei Province during the Chunyun period. The width of lines represents the migration volume from Wuhan to other cities. The pile graph represents the percentages of out-migration in Wuhan to Hubei province and the rest of China.
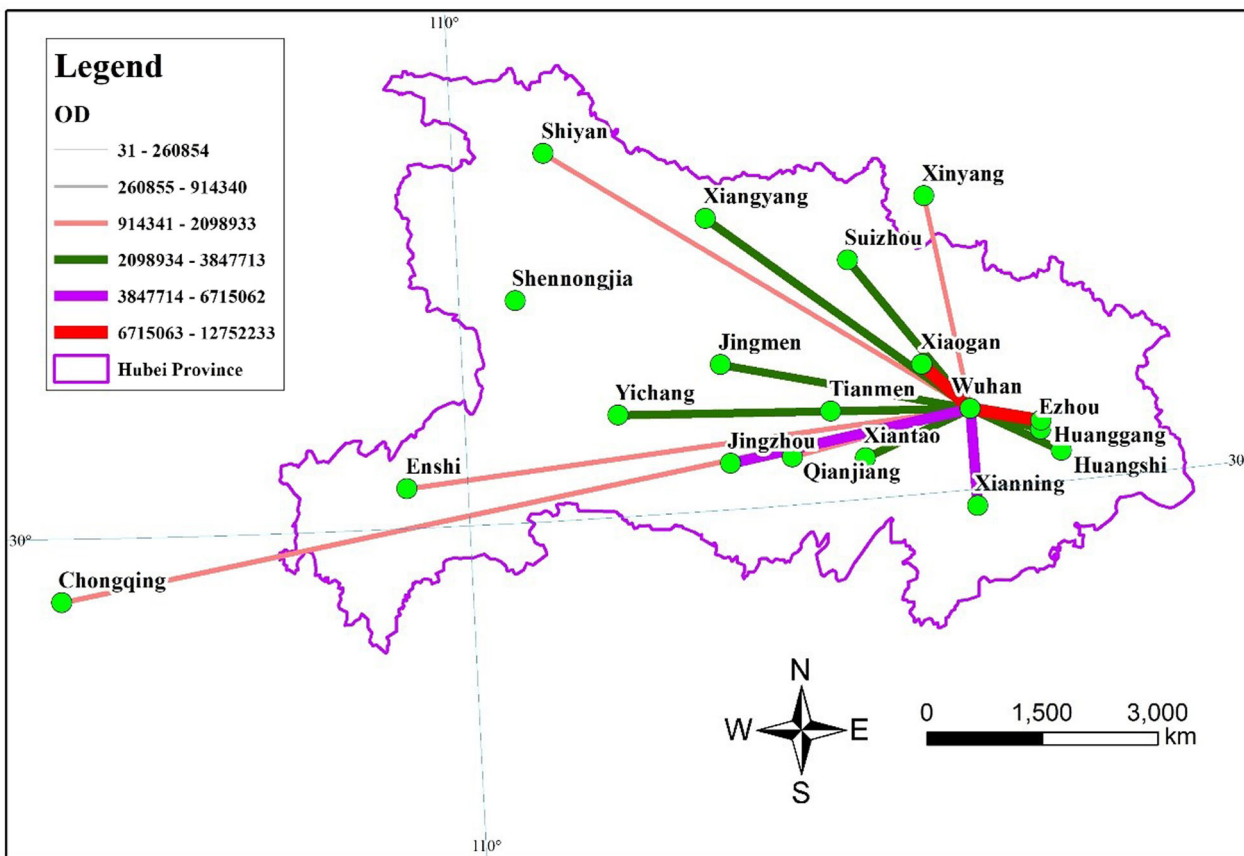
Fig. 9 The spatial distribution of OD from Wuhan (TOP 17) during the Chunyun period. The width of lines represents the migration volume from Wuhan to major cities.
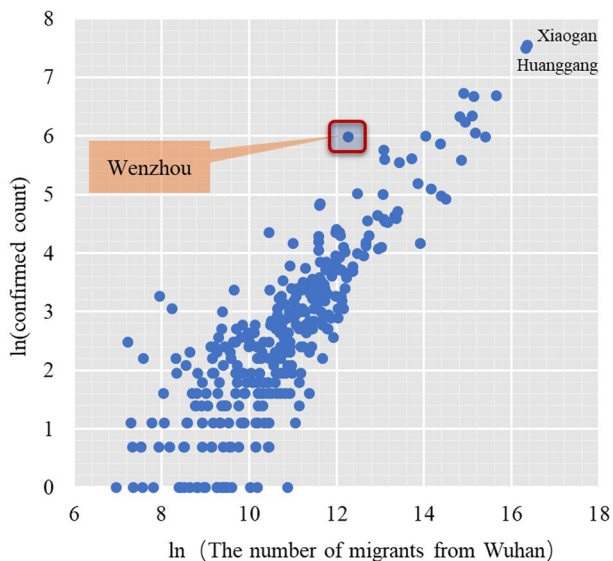
Fig. 10 The correlation between migration population from Wuhan (taken natural logarithm) and the number of covid-19 confirmed patients in other cities during the Chunyun period ($r = 0.943$, $p < 0.01$, $N = 321$). The city of Wenzhou is highlighted for being far from Wuhan and having large covid-19 cases.

national migration, epidemic control, and transportation policymaking.

Data availability

The data used to support the findings of this study have been deposited in the Zenodo repository (https://doi.org/10.5281/zenodo.3997309).

References

Bajardi P, Poletto C, Ramasco JJ, Tizzoni M, Colizza V, Vespignani A (2011) Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. PLoS ONE 6(1):e16591

Balcan D, Hu H, Goncalves B, Bajardi P, Poletto C et al. (2009) Seasonal transmission potential and activity peaks of the new influenza A(H1N1): a Monte Carlo likelihood analysis based on human mobility. BMC Med 7:45

Barabasi AL, Albert R (1999) Emergence of scaling in random networks. Science 286(5439):509–512

Bengtsson L, Gaudart J, Lu X, Moore S, Wetter E, Sallah K, Piarroux R (2015) Using mobile phone data to predict the spatial spread of cholera. Sci Rep 5:8923

Brooks CP, Antonovics J, Keitt TH (2008) Spatial and temporal heterogeneity explain disease dynamics in a spatially explicit network mode. Am Naturalist 172(2):149–159

Calabrese F, Pereira FC, Lorenzo G, Liu L, Ratti C (2010) The geography of taste: analyzing cell-phone mobility and social events. In: Pervasive Computing. Lecture Notes in Computer Science, Vol. 6030, Springer, Berlin, pp. 22–37

Chen G, Wang X, Li X (2012) Introduction to complex networks: models, structures and dynamics. Higher Education Press

Chung CKL, Xu J, Zhang M (2020) Geographies of Covid-19: how space and virus shape each other. Asian Geographer 37(2):99–116

Colizza V, Barrat A, Barthélemy M, Vespignani A (2006) The role of the airline transportation network in the prediction and predict-ability of global epidemics. Proc Natl Acad Sci USA 103:2015–2020

Deville P, Linard C, Martin S, Gilbert M, Stevens FR, Gaughan AE, Blodel VD, Tatem AJ (2014) Dynamic population mapping using mobile phone data. Proc Natl Acad Sci USA 11:15888–15893

Ding R, Ujang N, Hamid HB, Manan MSA, Li R, Albadareen SSM, Nochian A, Wu J (2019) Application of complex networks theory in urban traffic network researches. Netw Spatial Econ 19(4):1281–1317

Freeman LC (1977) A set of measures of centrality based on betweenness. Sociometry 40(1):35–41

Freeman LC, Douglas R, Mulholland RR (1979–1980) Centrality in social networks. Soc Netw 2(2):119–141

Gariazzo C, Pelliccioni A (2019) A multi-city urban population movement study using mobile phone traffic data. Appl Spatial Anal Policy 12(4):753–771

Gonzàlez MC, Hidalgo C, Baràbasi A (2008) Understanding individual human mobility pattern. Nature 453:779–782

Gu Z, Gu L, Eils R, Schlesner M, Brors B (2014) Circlize implements and enhances circular visualization in R. Bioinformatics 30(19):2811–2812

Guimera R, Mossa S, Turtschi A, Amaral LAN (2005) The worldwide air transportation network: anomalous centrality, subnetwork structure, and cities' global roles. Proc Natl Acad Sci USA 102(22):7794–7799

Hawelka B, Sitko I, Beinat E, Sobolevsky S, Kazakopoulos P, Ratti C (2014) Geolocated Twitter as proxy for global mobility patterns. Cartogr Geogr Inform Sci 41(3):260–71

Hu H (1935) The distribution of Chinese population. Acta Geogr Sinica 2:33–74

Hu Z et al. (2016) Spatiotemporal patterns of urban-rural development and transformation in east of the 'Hu Huanyong Line', China. ISPRS Int J Geo-Inform 5:24

Shumway JM, Otterstrom SM (2010) Spatial patterns of migration and income change in the mountain west: the dominance of service-based, amenity-rich counties. Profession Geogr 53(4):492–502

Kamath KYJ, Caverlee J, Cheng Z, Sui DZ (2012) Spatial influence vs. subnetwork influence: modeling the global spread of social media. In: Proceedings of the 21st ACM International Conference on Information and Knowledge Management. ACM, Maui, HI, pp. 962–971

Khan NU, Wan W, Yu S (2020) Spatiotemporal analysis of tourists and residents in Shanghai based on location-based social network's data from Weibo. ISPRS Int J Geo-Inf 9:70

Lee KS, You SY, Eom JK, Song J, Min JH (2018) Urban spatiotemporal analysis using mobile phone data: case study of medium- and large-sized Korean cities. Habitat Int 73(2018):6–15

Liu X, Wang M (2016) How polycentric is urban China and why? A case study of 318 cities. Landsc Urban Plan 151:10–20

Liu Y, Xu W, Shen J, Wang G (2016) Market expansion, state intervention and wage differentials between economic sectors in urban China: A multilevel analysis. Urban Stud 54(11):2631–2651

Liu Y, Xu W (2017) Destination choices of permanent and temporary migrants in China, 1985–2005. Population Space Place 23(1):1–17

Liu Z, Gu H (2019) Evolution characteristics of spatial concentration patterns of interprovincial population migration in China from 1985 to 2015. Applied Spatial Analysis and Policy

Louail T, Lenormand M, Ros OGC, Picornell M, Herranz R, Frias-Martinez E, Ramasco JJ, Barthelemy M (2014) From mobile phone data to the spatial structure of cities. Sci Rep 4:5276

Ma H, Fang C, Pang B, Wang S (2015) Structure of Chinese city network as driven by technological knowledge flows. Chinese Geogr Sci 25(4):498–510

Naaman M, Zhang AX, Brody S et al. (2012) On the study of diurnal urban routines on Twitter. In: 6th International AAAI Conference on Weblogs and Social Media (ICWSM-12). Dublin, Ireland, pp. 4–7

Newman MEJ (2001) Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. Phys Rev E 64(2):016132

Pei T, Sobolevsky S, Ratti C, Shaw S-L, Li T, Zhou C (2014) A new insight into land use classification based on aggregated mobile phone data. Int J Geogr Inform Sci 28(9):1988–2007

Picornell M, Ruiz T, Lenormand M, Ramasco JJ, Dubernet T, Frías-Martínez E (2015) Exploring the potential of phone call data to characterize the relationship between social network and travel behavior. Transportation 42(4):647–68

Qi W, Abel G, Muttarak R, Liu S (2017) Circular visualization of China's internal migration flows 2010–2015. Environ Plan A 49(11):2432–2436

Ravenstein E (1884) The laws of migration. J Stat Soc London 48

Read JM et al. (2020) Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. medRxiv 2020.2001.2023.20018549

Rinzivillo S, Mainardi S, Pezzoni F, Coscia M, Pedreschi D, Giannotti F (2012) Discovering the geographical borders of human mobility. KI-Künstliche Intelligenz 26(3):253–260

Sagl G, Delmelle E, Delmelle E (2014) Mapping collective human activity in an urban environment based on mobile phone data. Cartography Geogr Inform Sci 41(3):272–285

Shanmukhappa T, Ho IW-H, Tse CK (2018) Spatial analysis of bus transport networks using network theory. Phys A 502(C):295–314

Shen J (2020) Analyzing the determinants of the spread of covid-19 among the provincial regions in China. Preprint.

Wang L (2018) High-speed rail services development and regional accessibility restructuring: the case of the Yangtze River Delta, China. Transport Policy 72:34–44

Wang L, Acheampong RA, He S (2020) High-speed rail network development effects on the growth and spatial dynamics of knowledge-intensive economy in major cities of China. Cities 105:102772

Wang Y, Deng Y, Ren F, Zhu R, Wang P, Du T, Du Q (2020) Analysing the spatial configuration of urban bus networks based on the geospatial network analysis method. Cities 96:102406

Watts DJ, Strongatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393(6684):440–442

Wei S, Yuan J, Qiu Y, Luan X, Han S, Zhou W et al. (2017) Exploring the potential of open big data from ticketing websites to characterize travel patterns within the Chinese high-speed rail system. PLoS ONE 12(6):e0178023

Wei S, Teng SN, Li HJ, Xu J, Ma H, Luan XL et al. (2019) Hierarchical structure in the world's largest high-speed rail network. PLoS ONE 14(2):e0211052

Wei S, Xu J, Ma H (2019) Exploring public bicycle network structure based on complex network theory and shortest path analysis: The public bicycle system in Yixing, China. Transport Plan Technol 42(3):293–307

Widhalm P, Yang Y, Ulm M, Athavale S, González MC (2015) Discovering urban activity patterns in cell phone data. Transportation 42(4):597–623

Yin J et al. (2017) Depicting urban boundaries from a mobility network of spatial interactions: a case study of Great Britain with geo-located Twitter data. Int J Geogr Inform Sci 31(7):1293–1313

Acknowledgements

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.W.